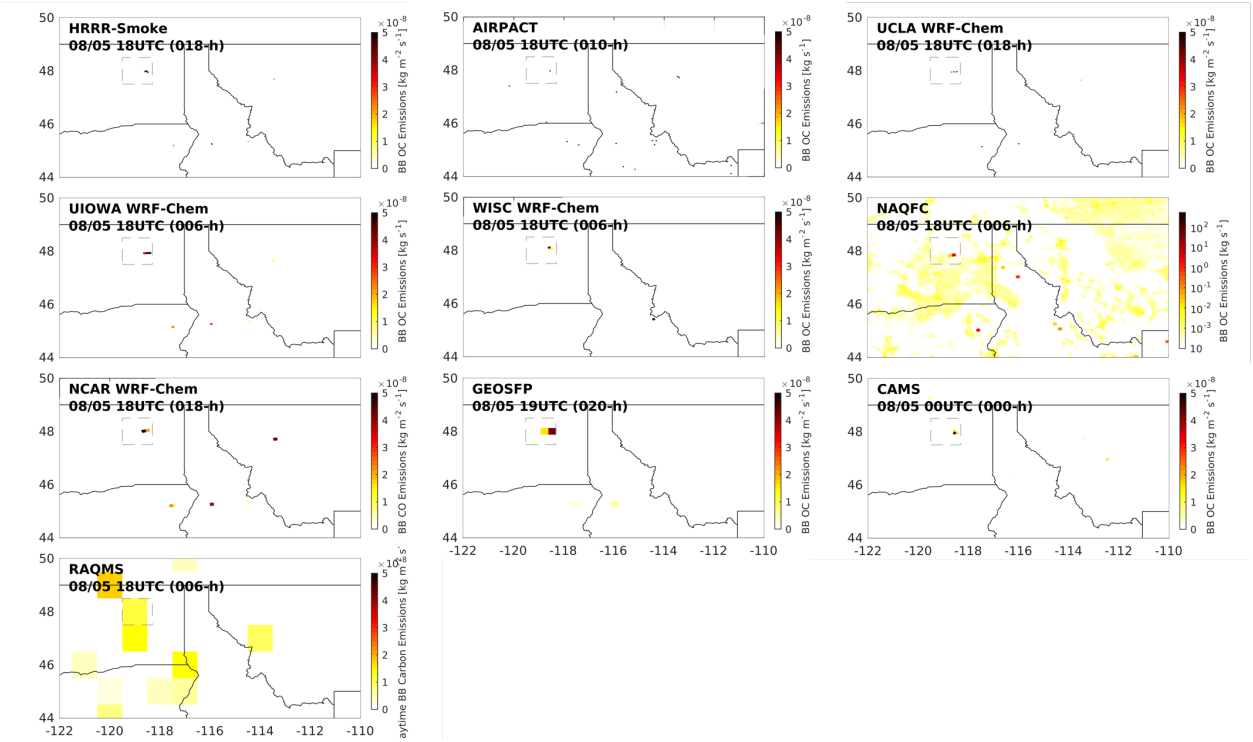


Figure S1. Map distributions of predicted biomass burning emissions used in models. The model name and valid time are labeled respectively on each panel with the lead time in parenthesis, e.g. 018-h. Note that owing to the availability of emission data, not all models incorporated in this work are shown. The dashed box represents the domain from which the emissions from the Williams Flats fire are extracted.

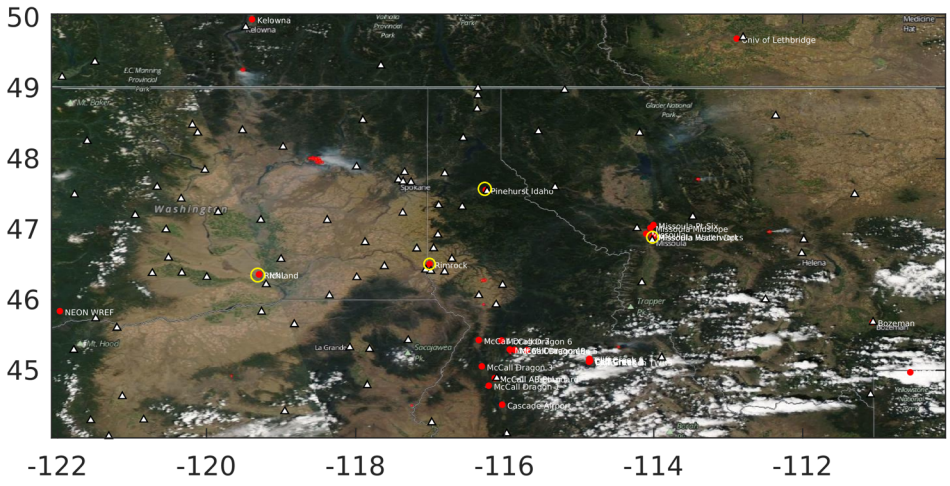


Figure S2. Locations of the AERONET and AENONET-DRAGON sites that operated during the FIREX-AQ field campaign, overlaid on the EOS-Aqua MODIS visible image on 5 August 2019 obtained from NASA Worldview (https://worldview.earthdata.nasa.gov/). The white triangles are surface monitoring stations in AirNow network.

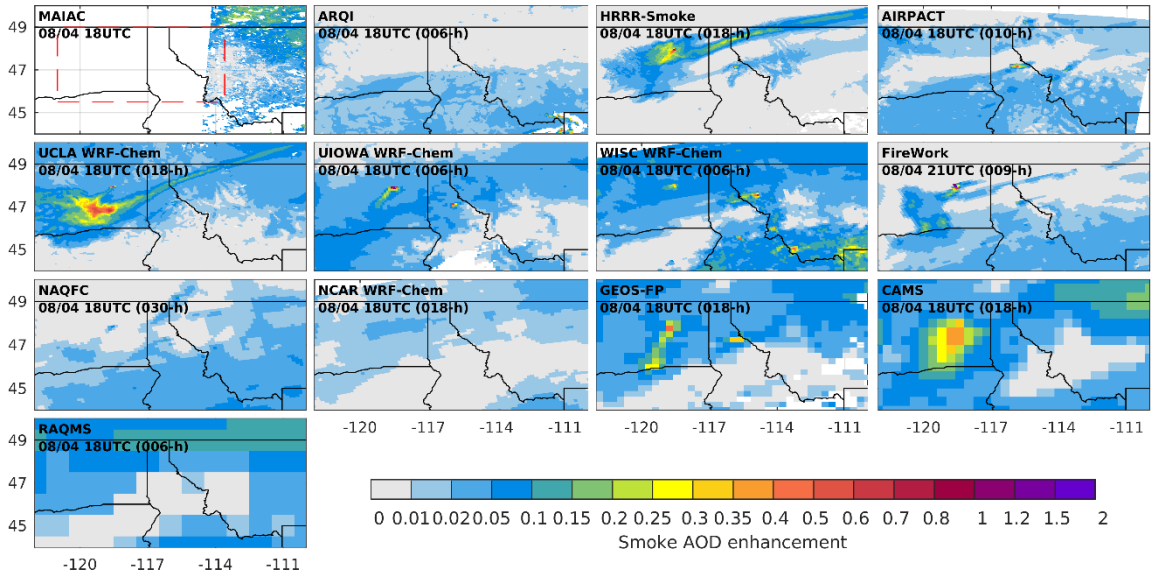


Figure S3. Map distributions of smoke AOD enhancements (sAOD) observed by MODIS MAIAC AOD data and the model forecasts for 1800 UTC 4 August 2019. The valid time of forecast results is shown in each panel with the lead time in parenthesis, e.g. 008-h. Background AOD has been subtracted, which is represented by the average of the lowest 20% values within the respective maps. Note that the areas in white are missing data that are related to cloud contaminations and masked according to quality filters (for the observations) or total column cloud water/cloud fractions (for the model forecasts). The dashed box in red in the observation map shows the area of interest in the verification of sAOD magnitude and spatial extent of the smoke plumes.

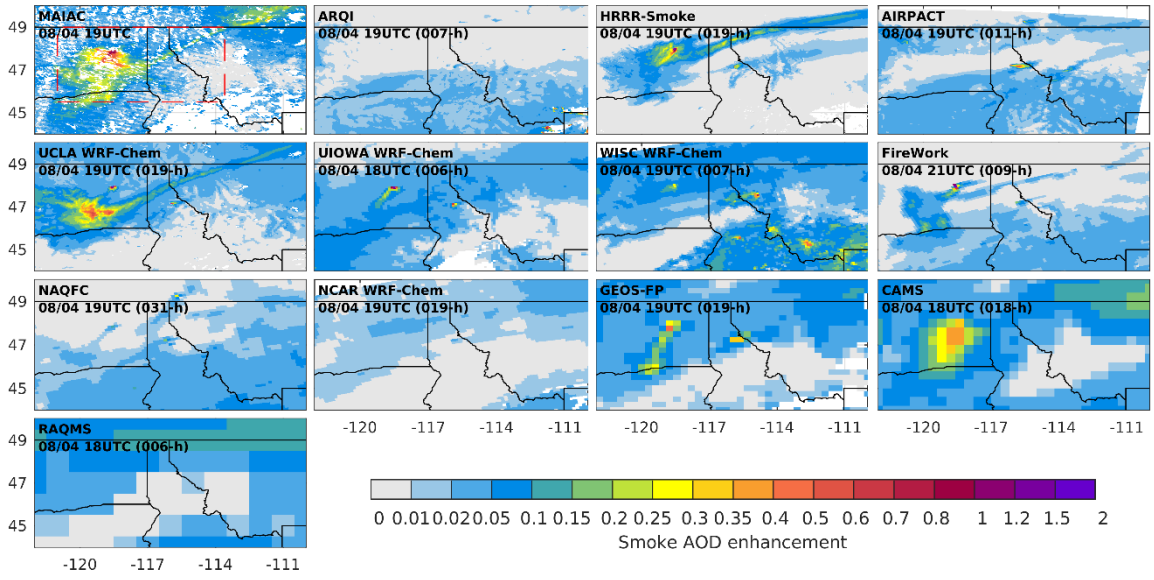


Figure S4. Same as Figure S3, but for 1900 UTC 4 August 2019.

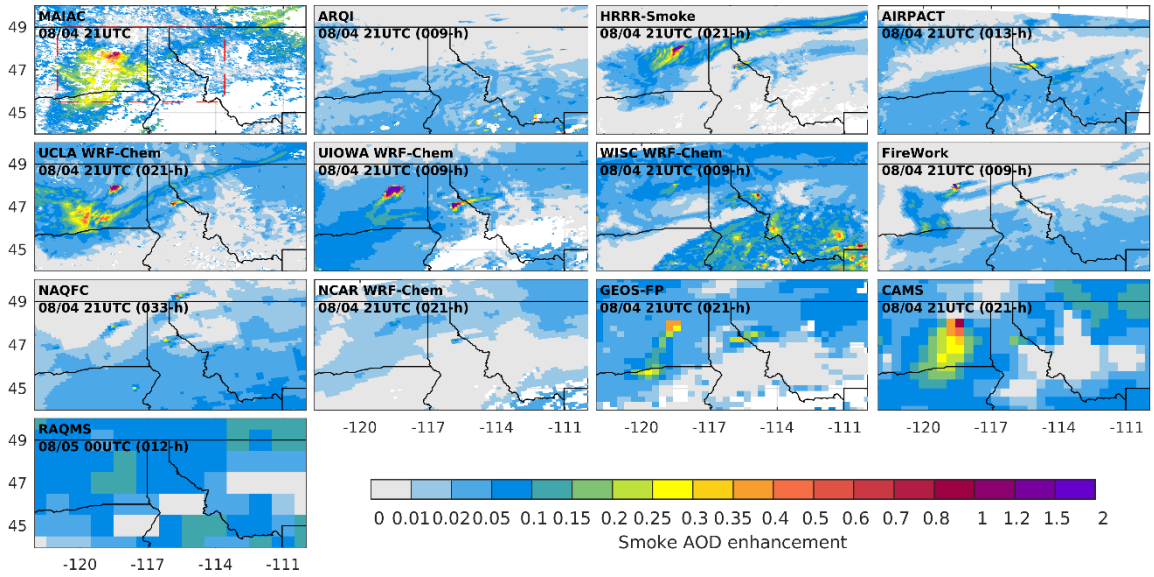


Figure S5. Same as Figure S3, but for 2100 UTC 4 August 2019.

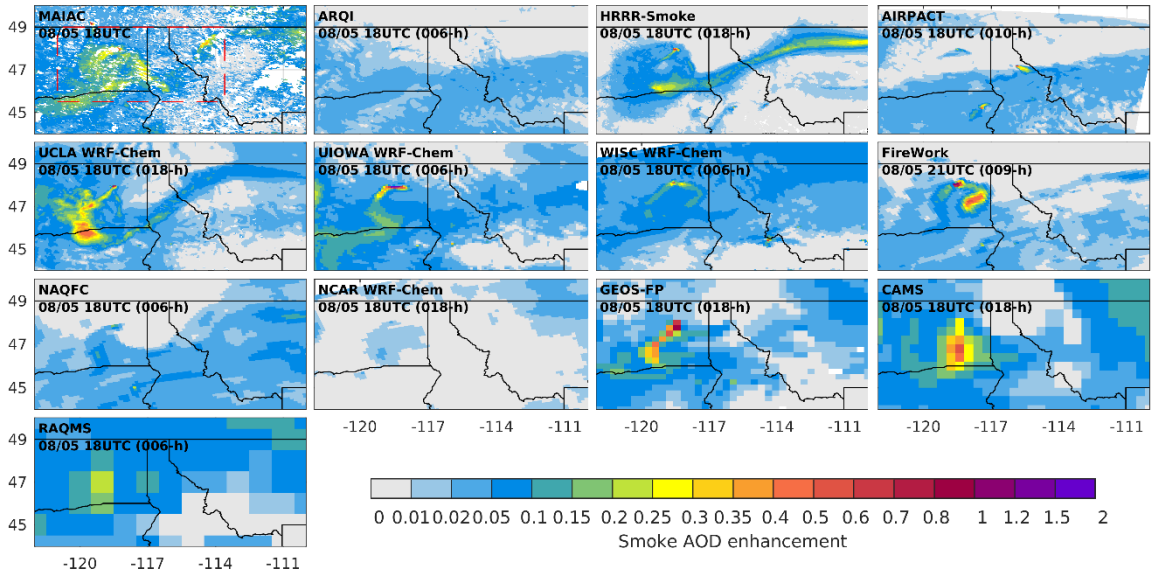


Figure S6. Same as Figure S3, but for 1800 UTC 5 August 2019.

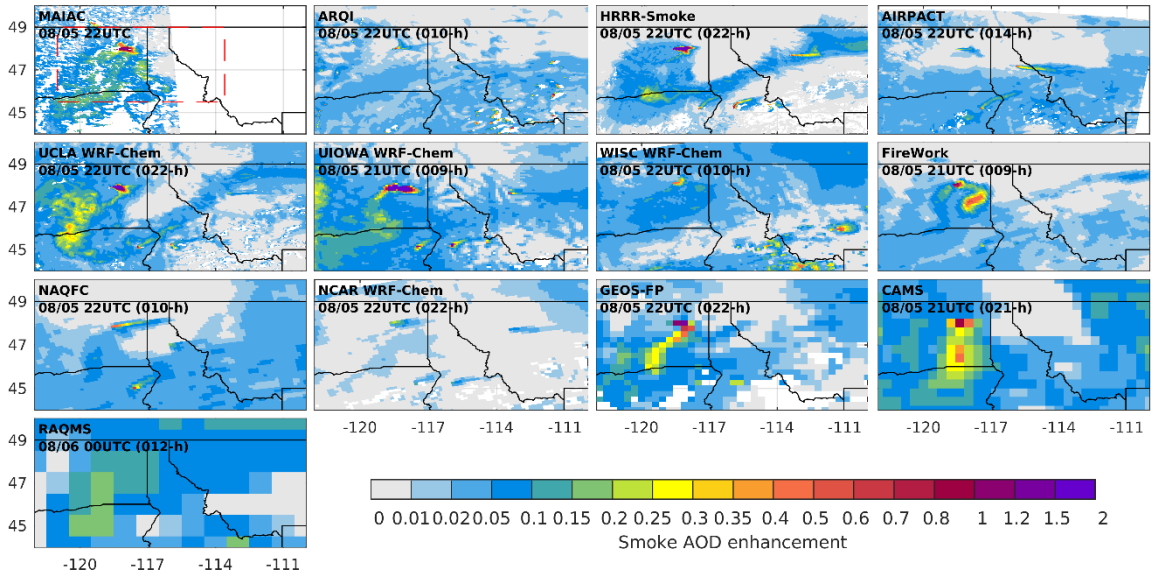


Figure S7. Same as Figure S3, but for 2200 UTC 5 August 2019.

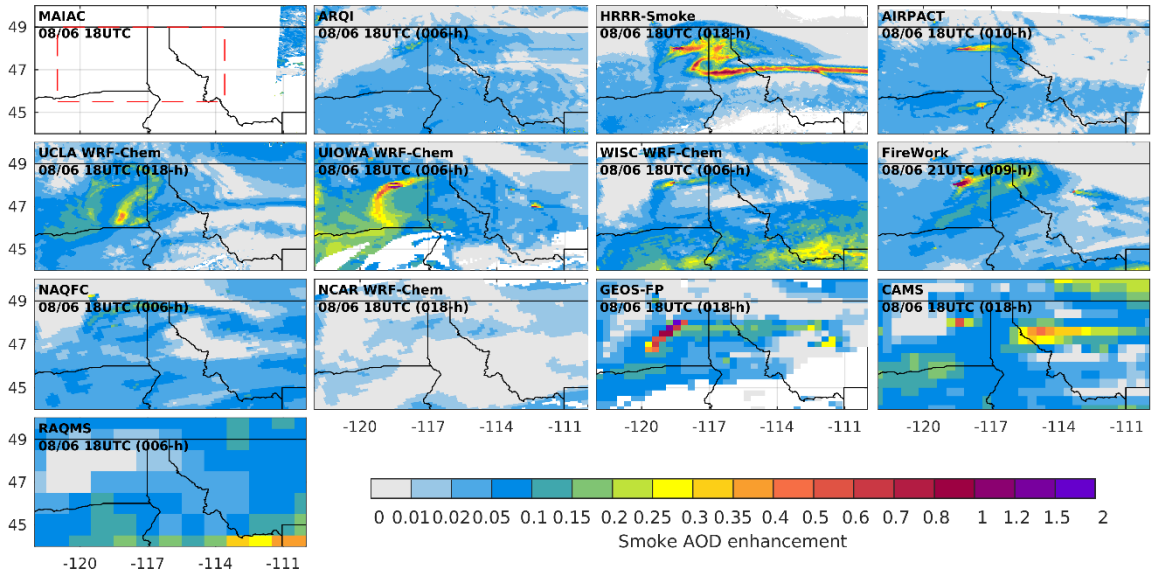


Figure S8. Same as Figure S3, but for 1800 UTC 6 August 2019.

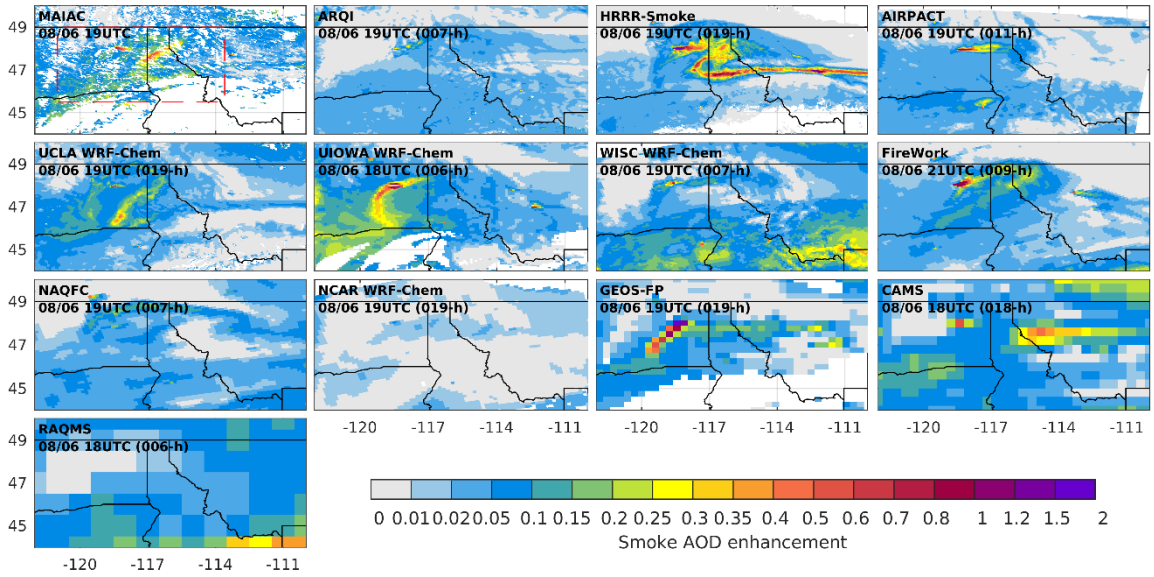


Figure S9. Same as Figure S3, but for 1900 UTC 6 August 2019.

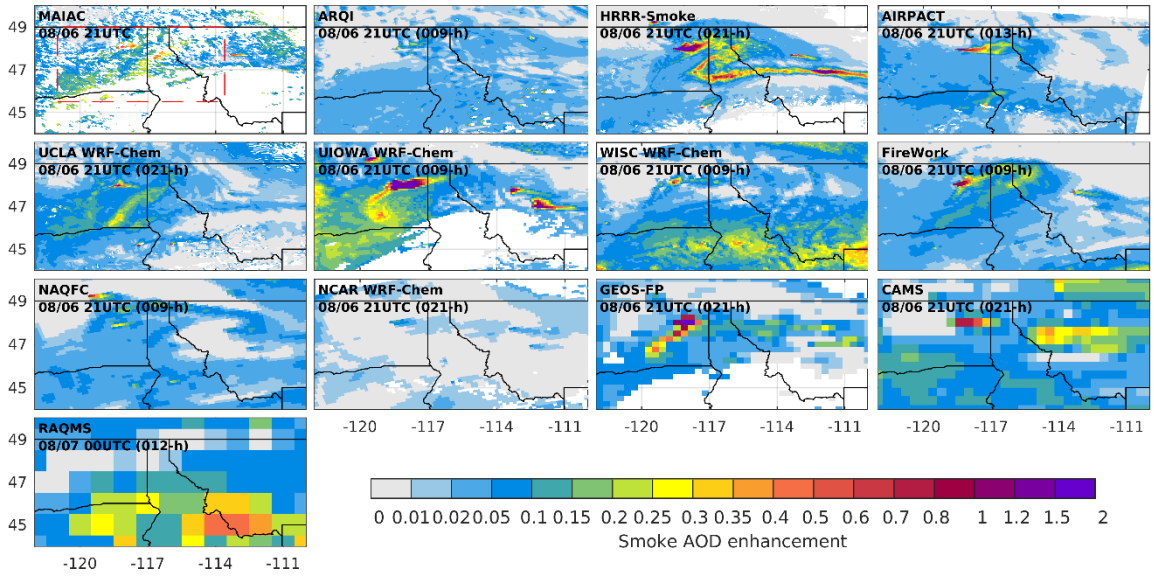


Figure S10. Same as Figure S3, but for 2100 UTC 6 August 2019.

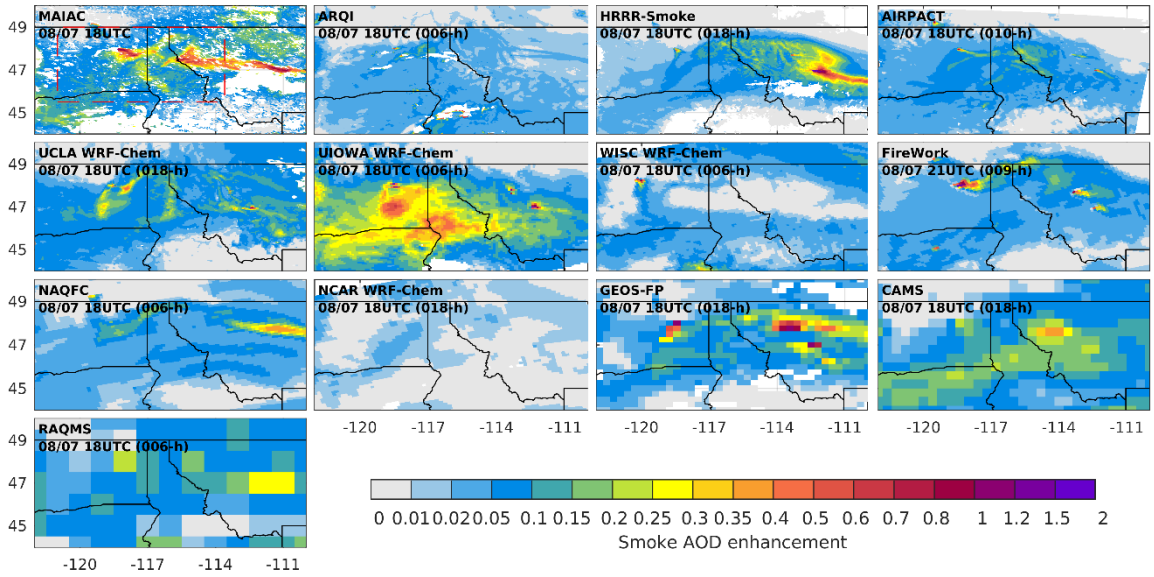


Figure S11. Same as Figure S3, but for 1800 UTC 7 August 2019.

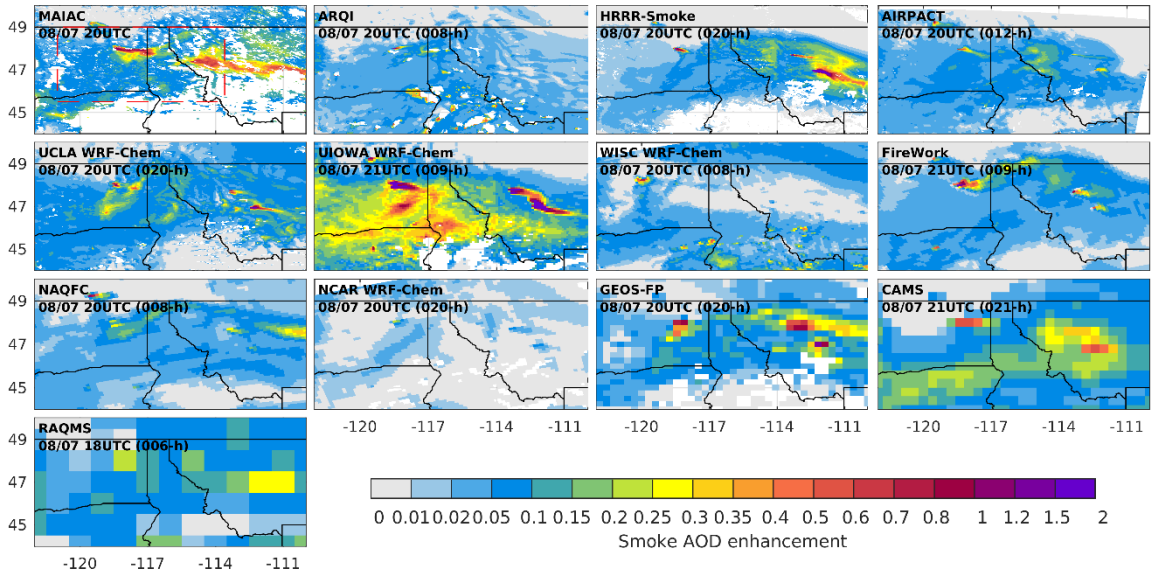


Figure S12. Same as Figure S3, but for 2000 UTC 7 August 2019.

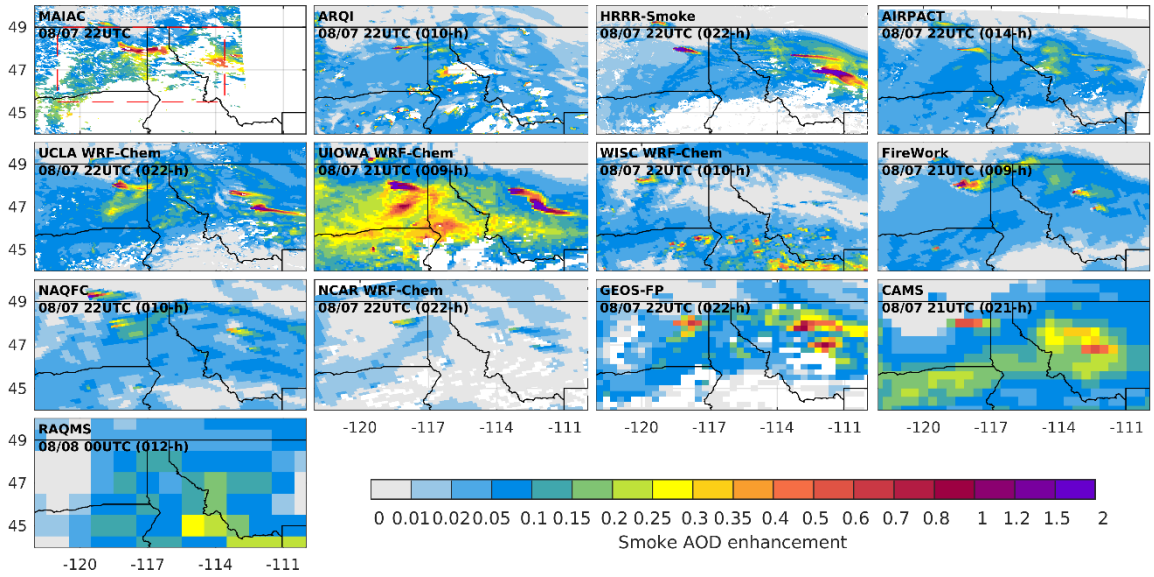


Figure S13. Same as Figure S3, but for 2200 UTC 7 August 2019.

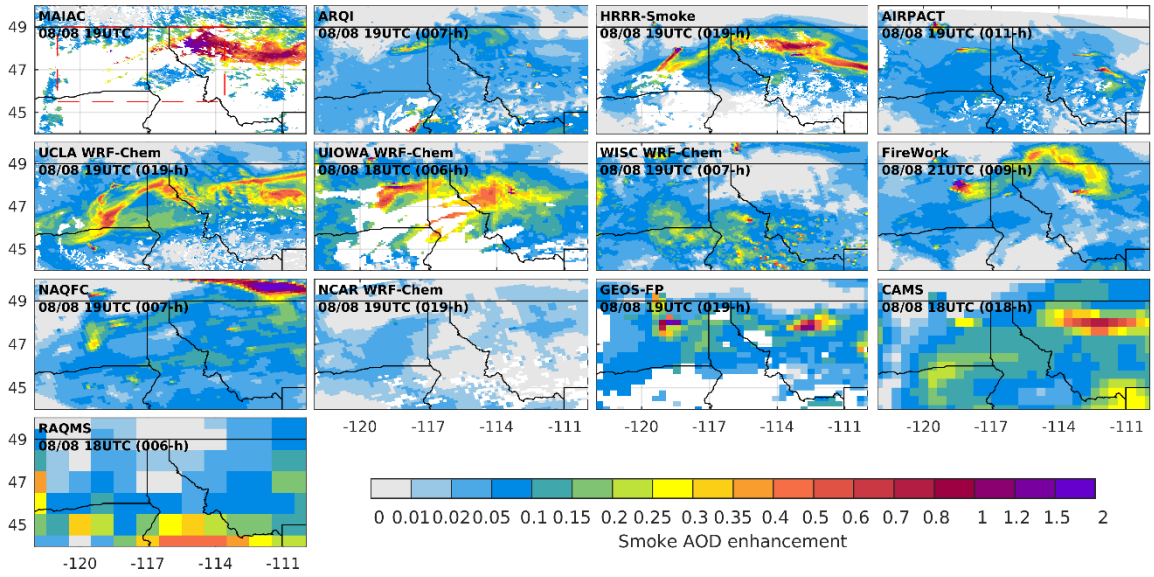


Figure S14. Same as Figure S3, but for 1900 UTC 8 August 2019.

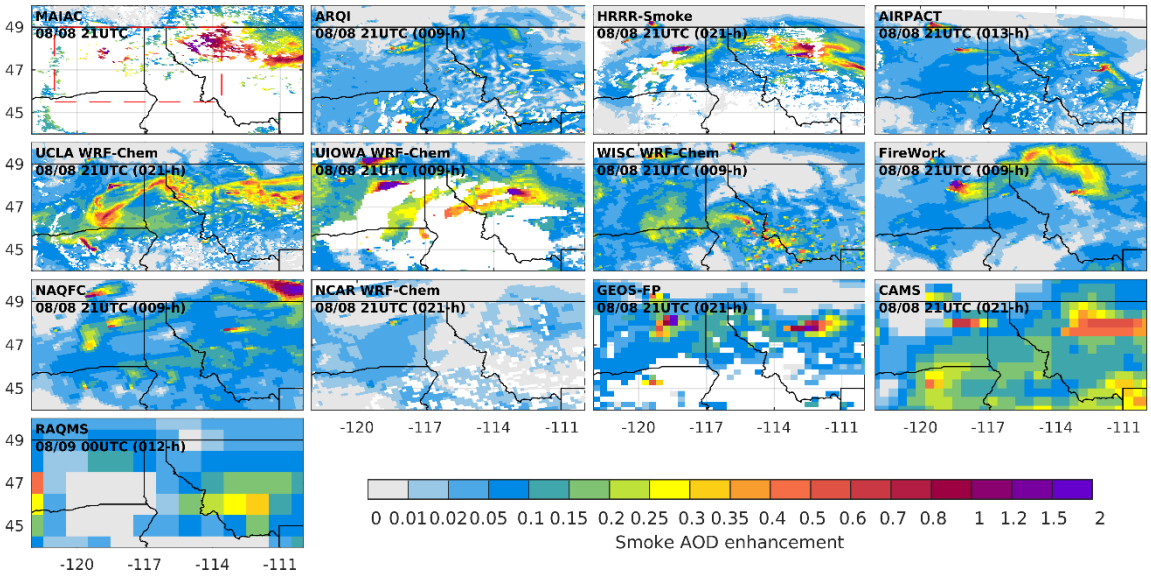


Figure S15. Same as Figure S3, but for 2100 UTC 8 August 2019.

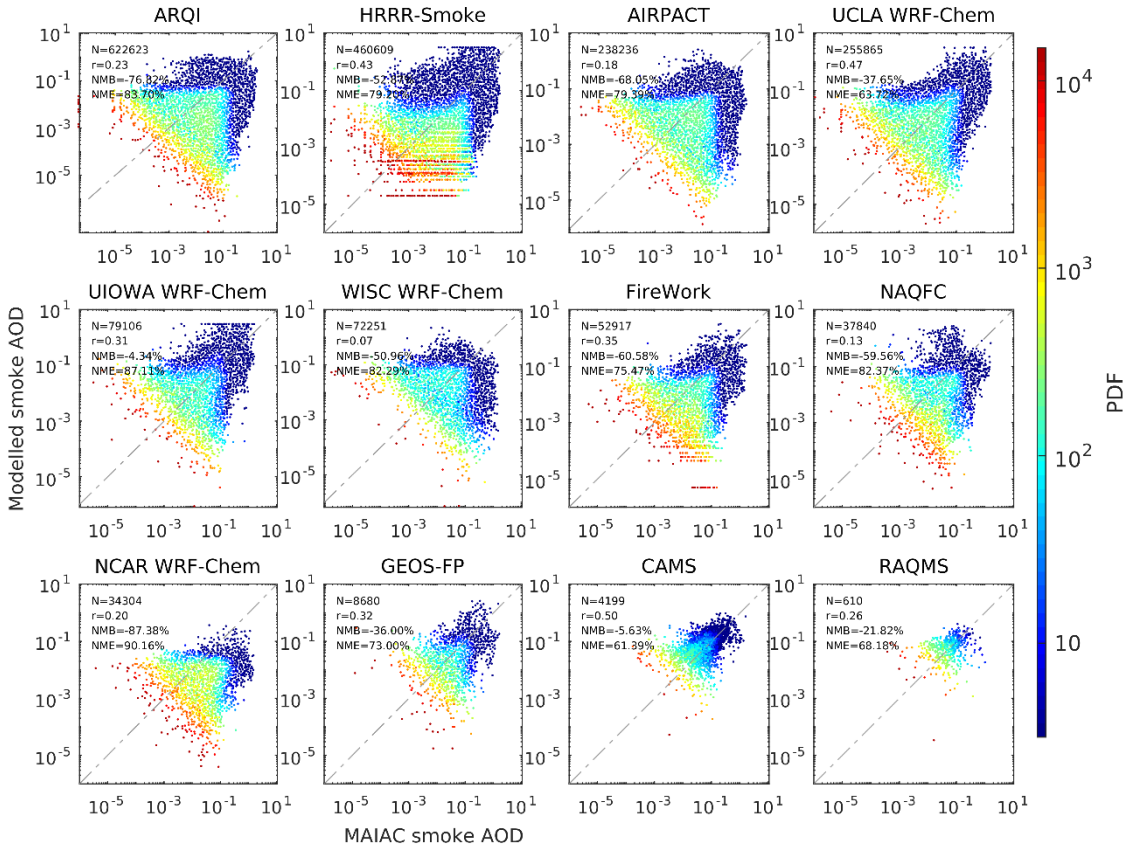


Figure S16. Scatter plots of predicted and observed smoke AOD enhancements (sAOD) on 4-8 August 2019. The data points are colored by two-dimensional probability density function (in log scale). The grey line represents 1:1. The

number of data pairs (N), correlation coefficient (r), Normalized Mean Error (NME), and Normalized Mean Bias (NMB) are labeled in each panel.

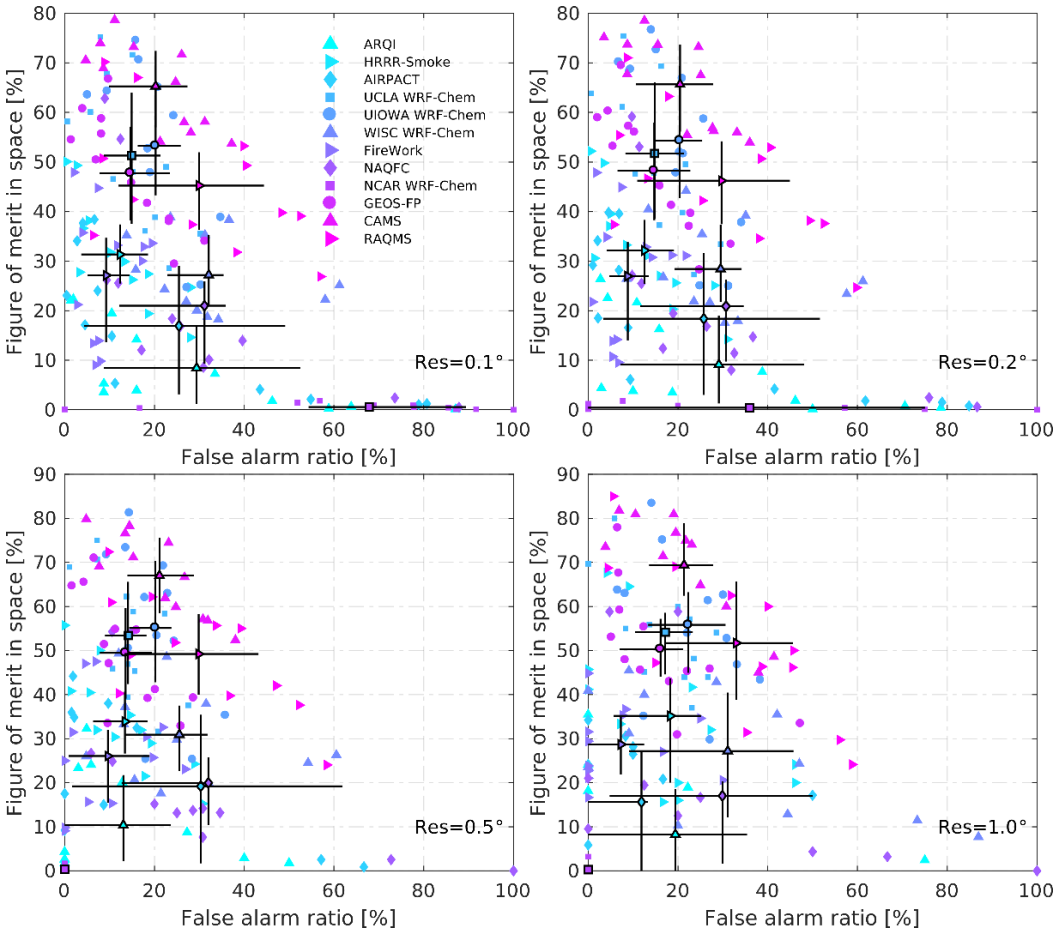


Figure S17. FMS and FAR scores for fire smoke AOD exceedance events (sAOD>0.05) forecasted by models compared against MODIS MAIAC AOD retrievals per hourly snapshot during 4-8 August 2019. The scores are derived using re-gridded data (resolutions as annotated). For each model, the markers with black edges represent median values and the horizontal and vertical black bars are the 25th to 75th percentiles.

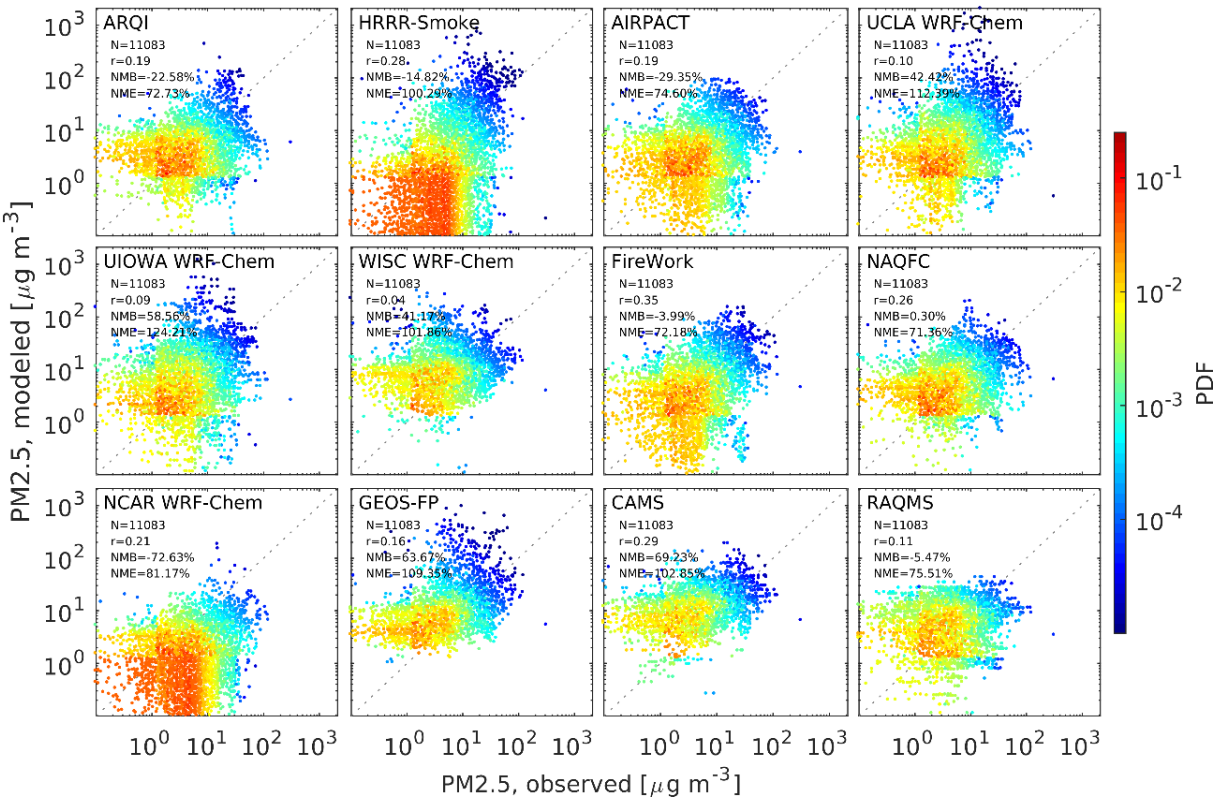


Figure S18. Similar as Figure S15, but for PM2.5 enhancements on 4-9 August 2019 for all the 83 monitoring stations.

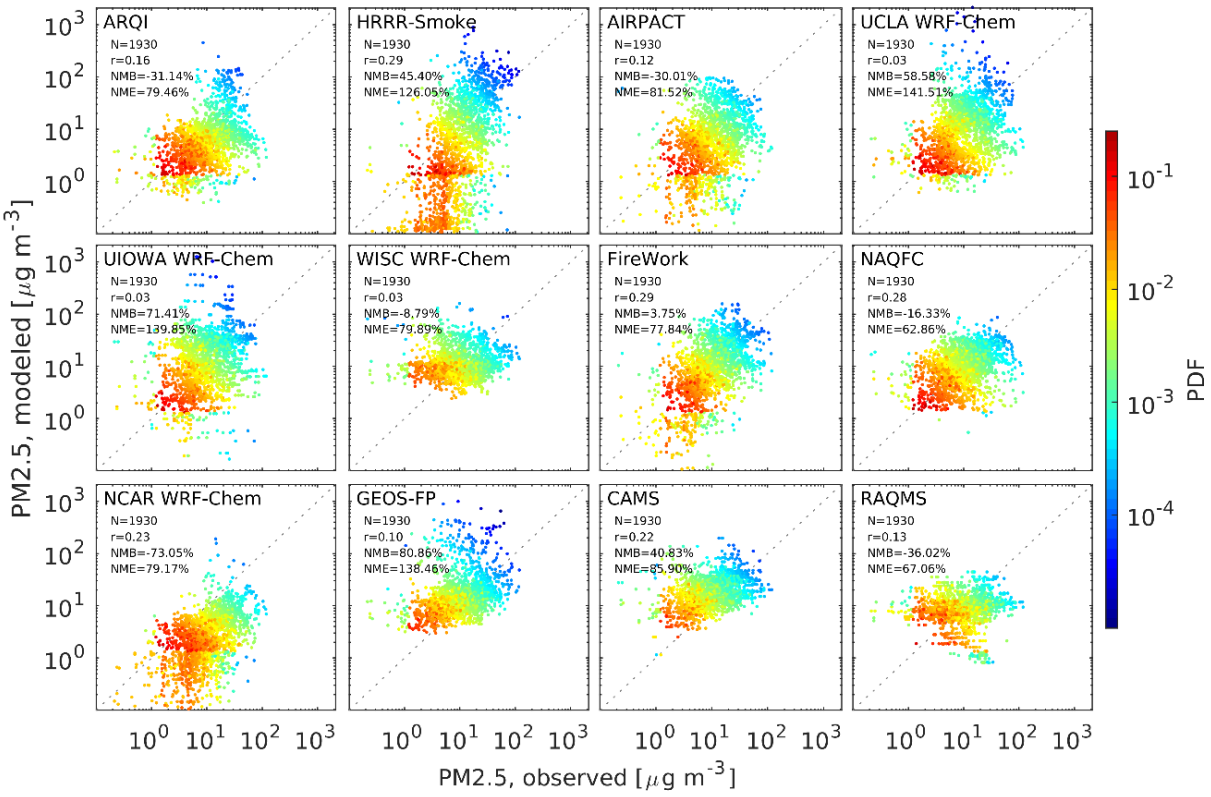


Figure S19. Similar as Figure S15, but for PM2.5 enhancements on 4-9 August 2019 for the 14 monitoring stations that represent impacts from fresh smoke plumes, i.e. classified as “Fp” stations.

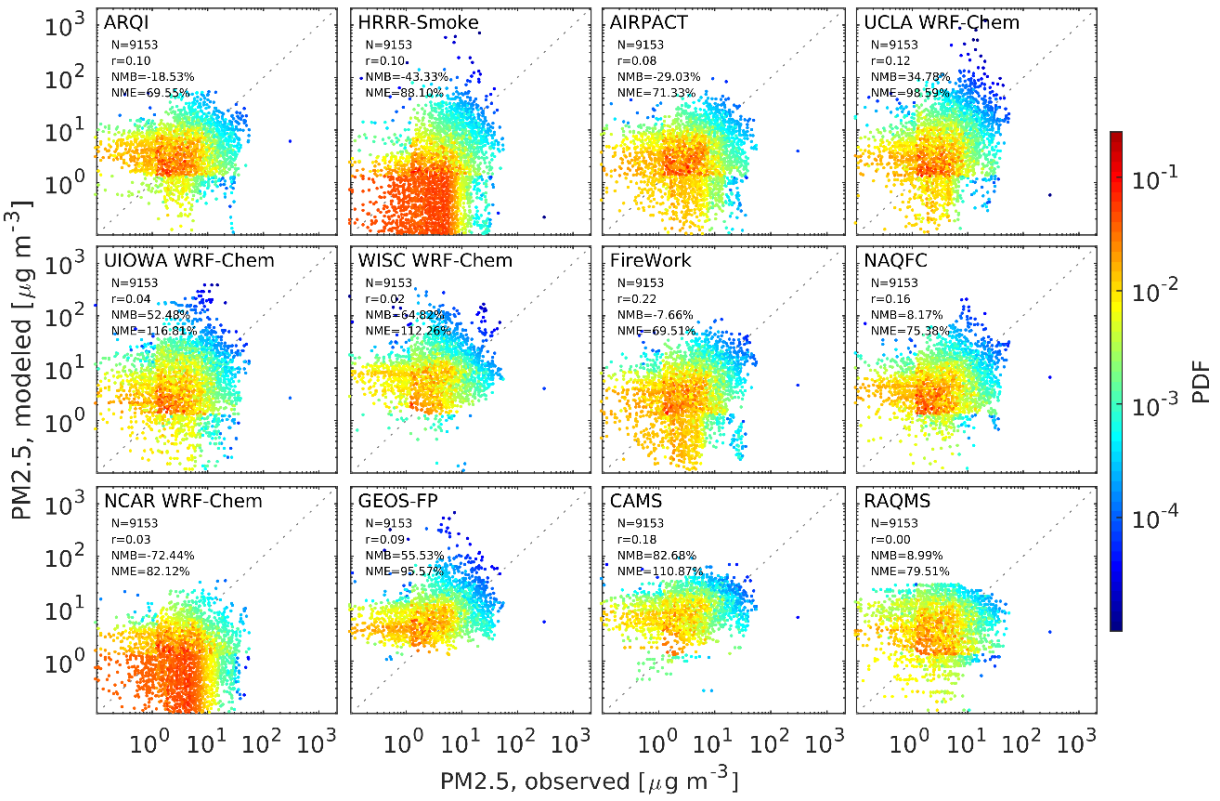


Figure S20. Similar as Figure S15, but for PM2.5 enhancements on 4-9 August 2019 for all the 69 monitoring stations that are classified as “Ot” stations.

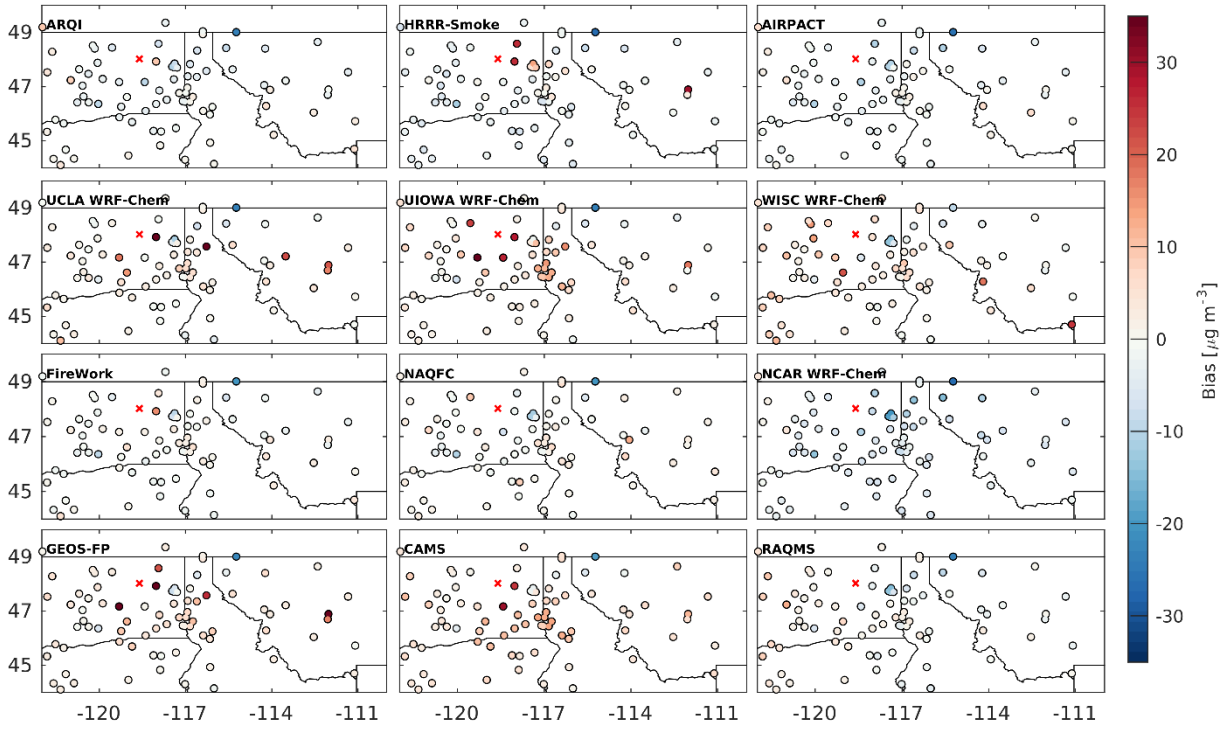


Figure S21. Map distributions of PM2.5 enhancement biases for the twelve models during 4-9 August 2019. The red multiplication sign represents the location of the Williams Flats Fire.

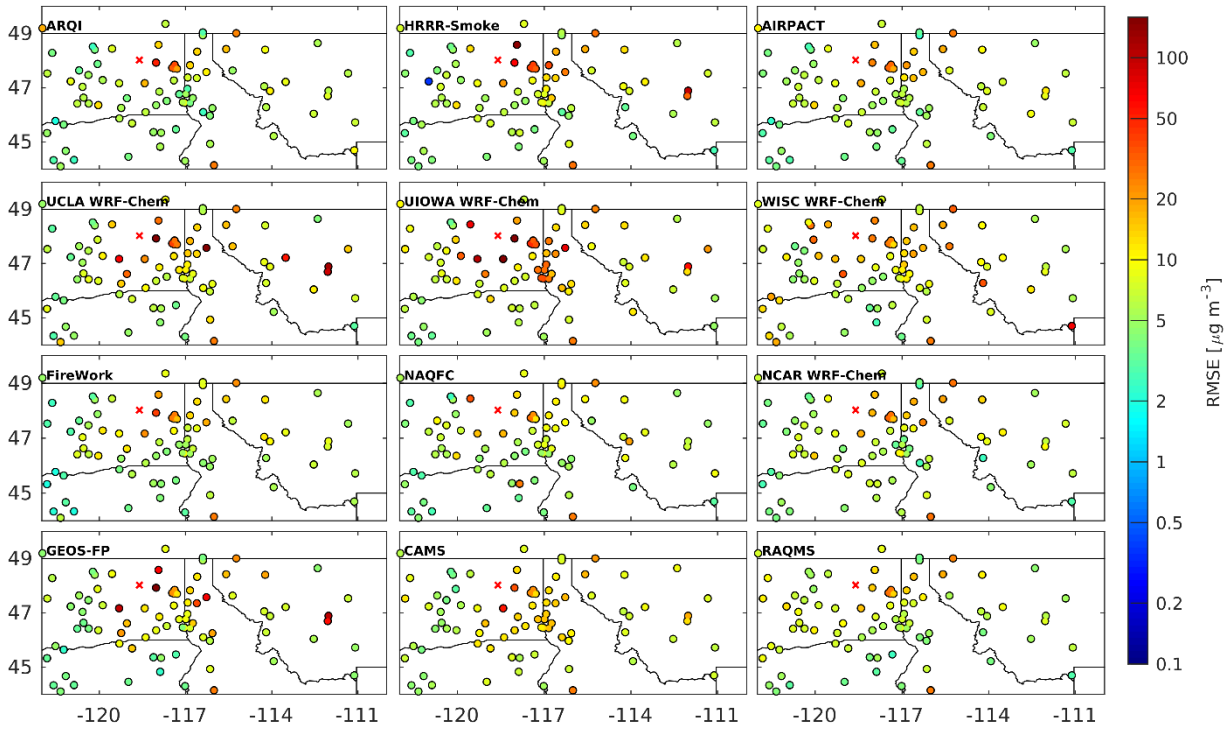


Figure S22. Similar as Figure S21, but for RMSE.

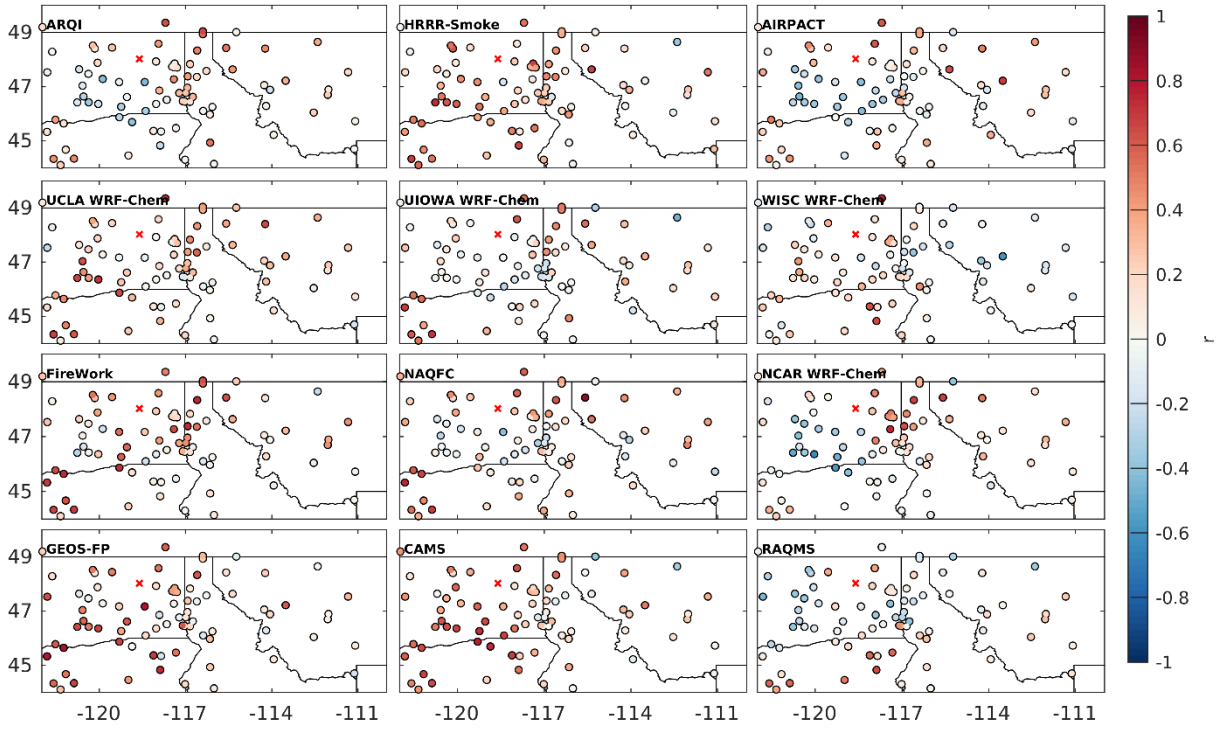


Figure S23. Similar as Figure S21, but for correlation coefficient (*r*).

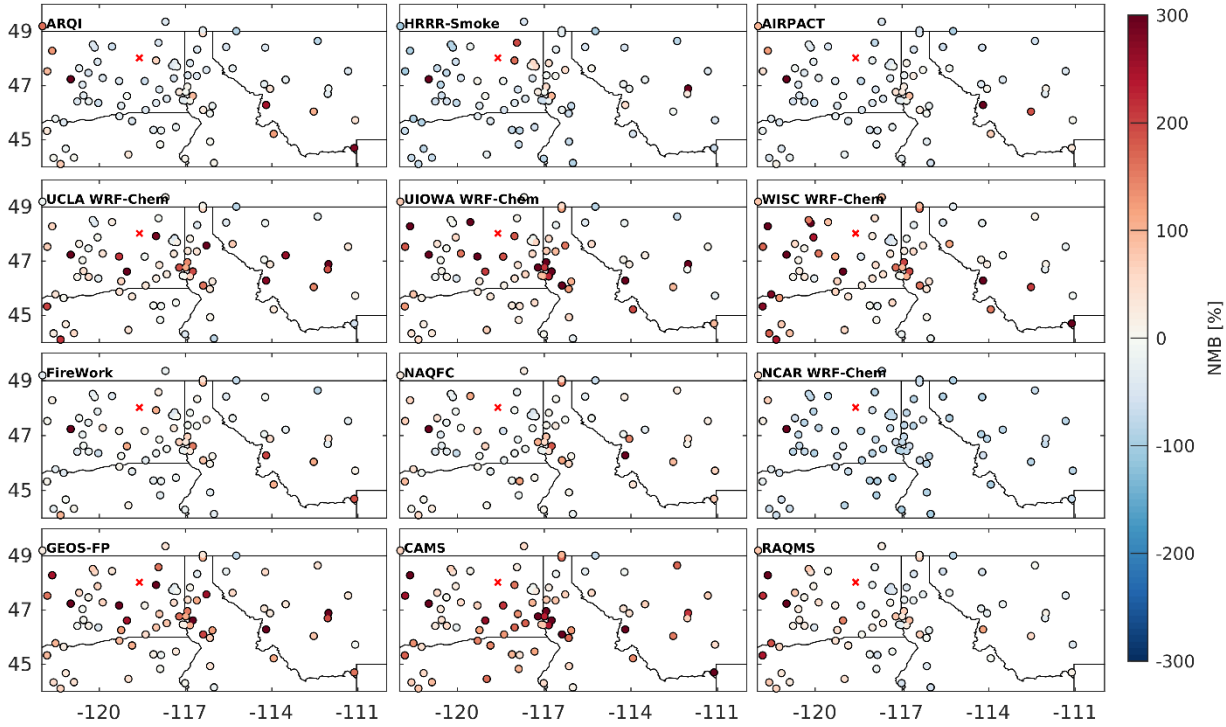


Figure S24. Similar as Figure S21, but for NMB.

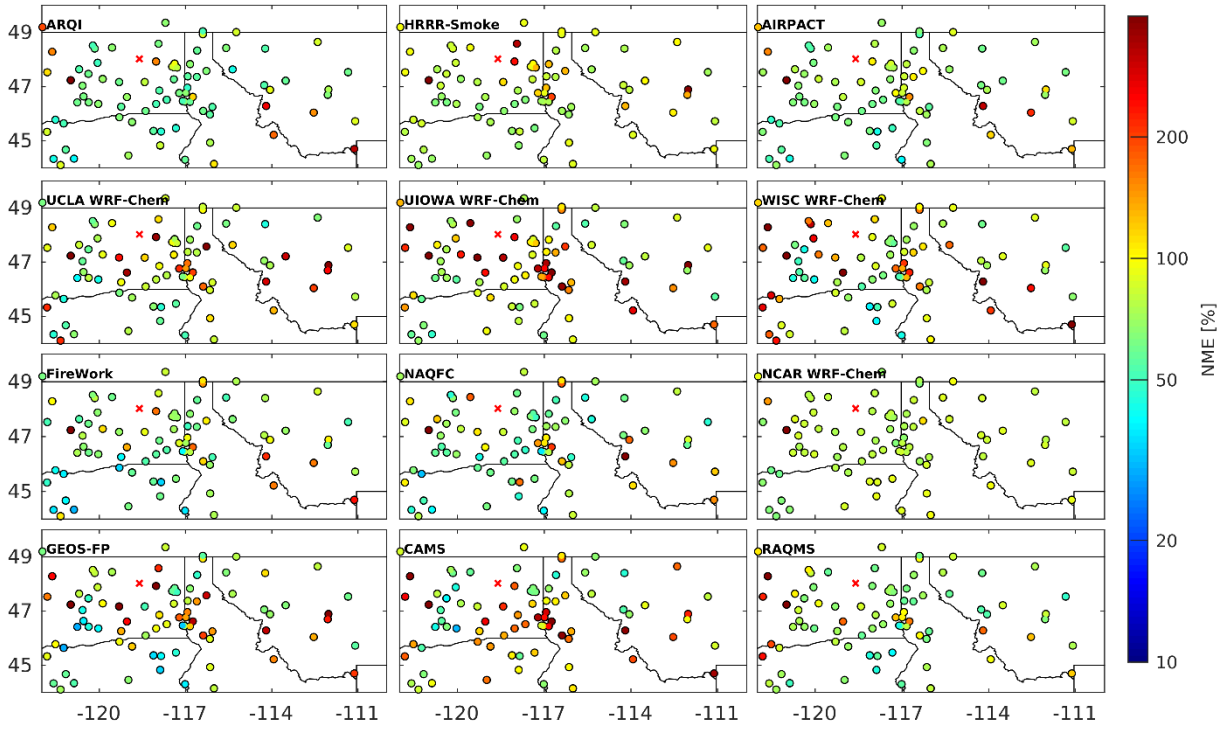


Figure S25. Similar as Figure S21, but for NME.

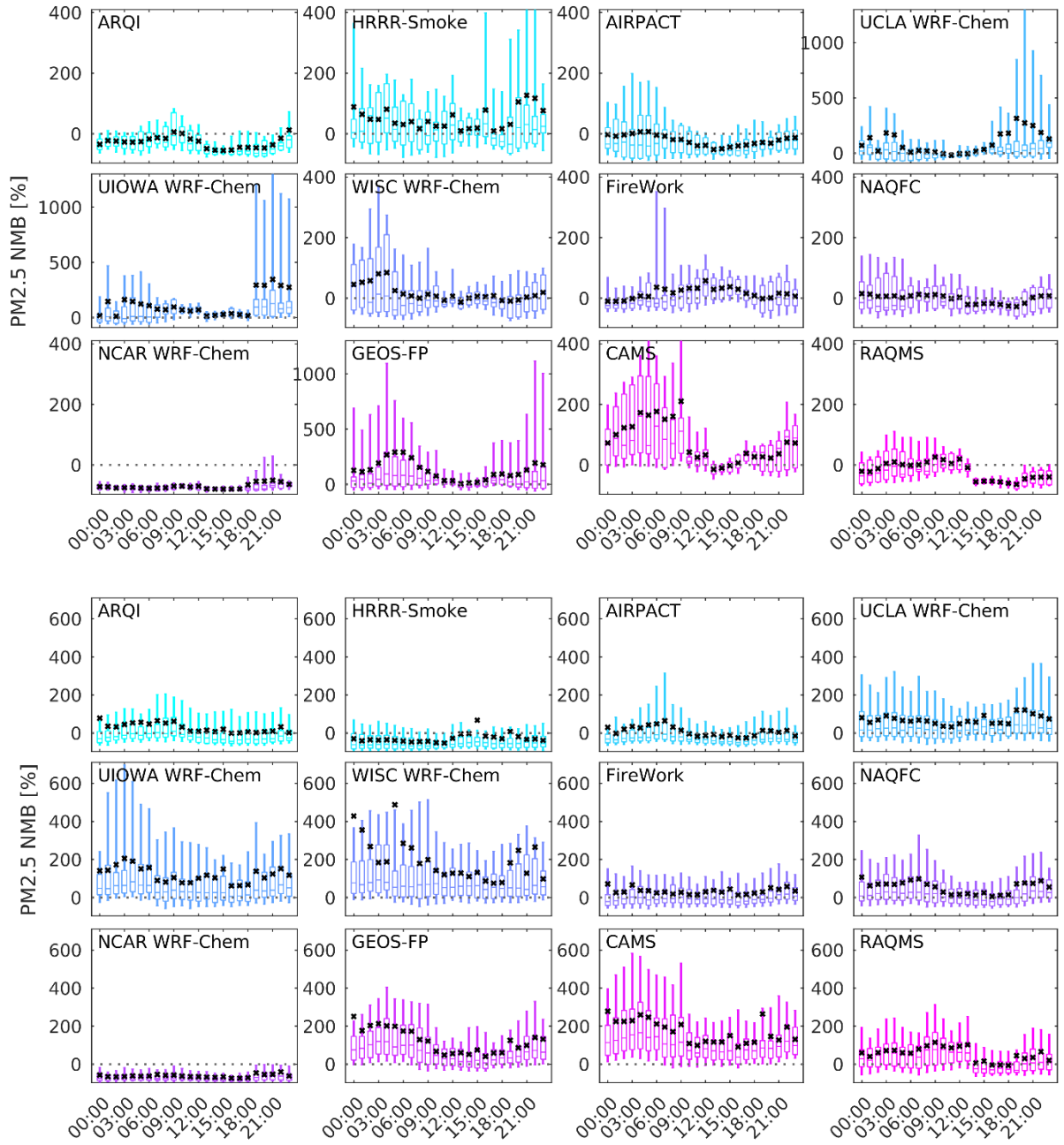


Figure S26. Box plots of NMB of PM2.5 enhancements forecast by the twelve models, compared against surface observations in each hour. Results are shown for two categories of sites, i.e. fresh-plume sites (Fp, upper panels) and other sites (Ot, lower panels). For each box, the central mark represents the median, and the bottom and top edges of box indicate the 25th and 75th percentiles, respectively. The whiskbn ers extend to the 10th and 90th percentiles, and the black "x" represents the average value.

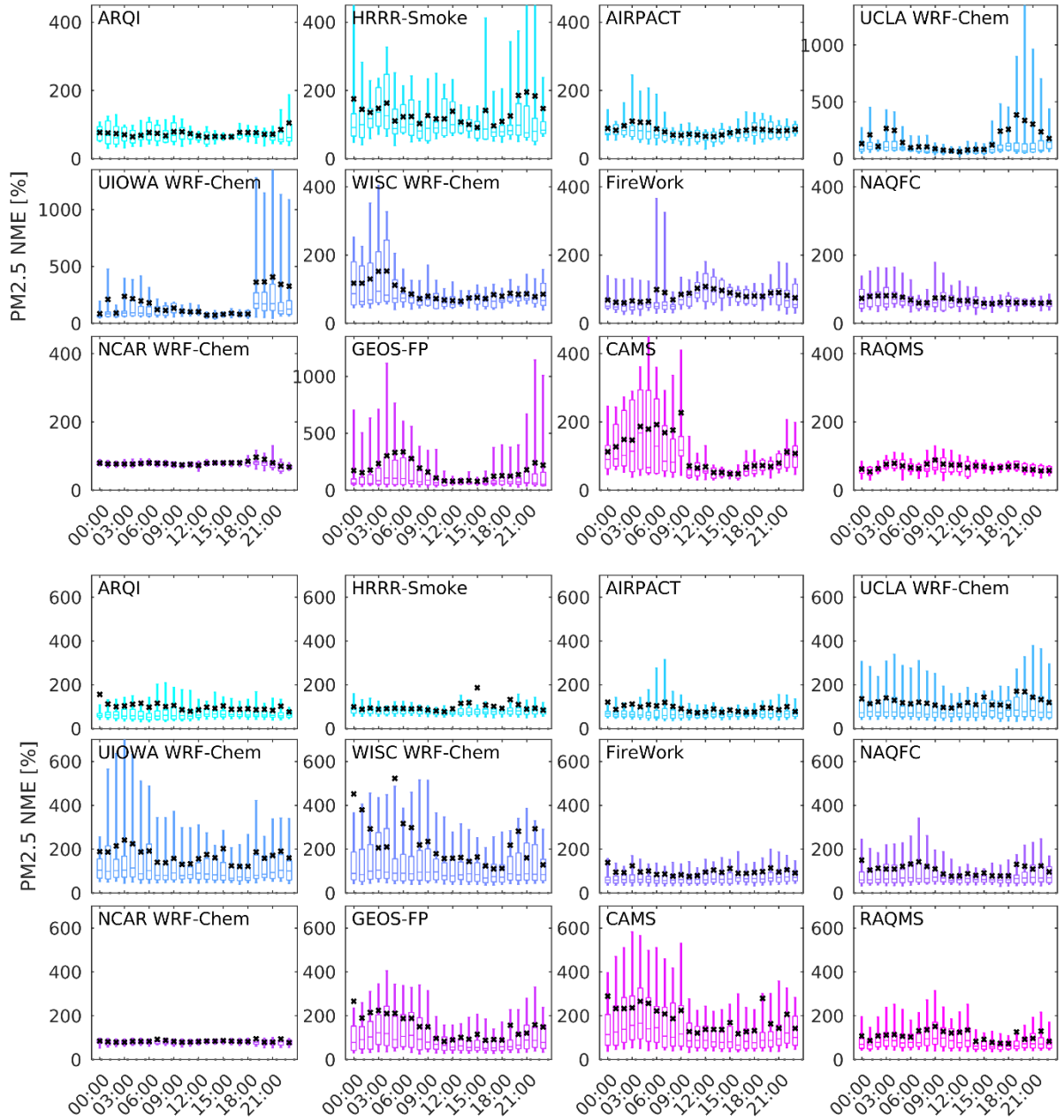


Figure S27. Similar as Figure S26, but for NME.

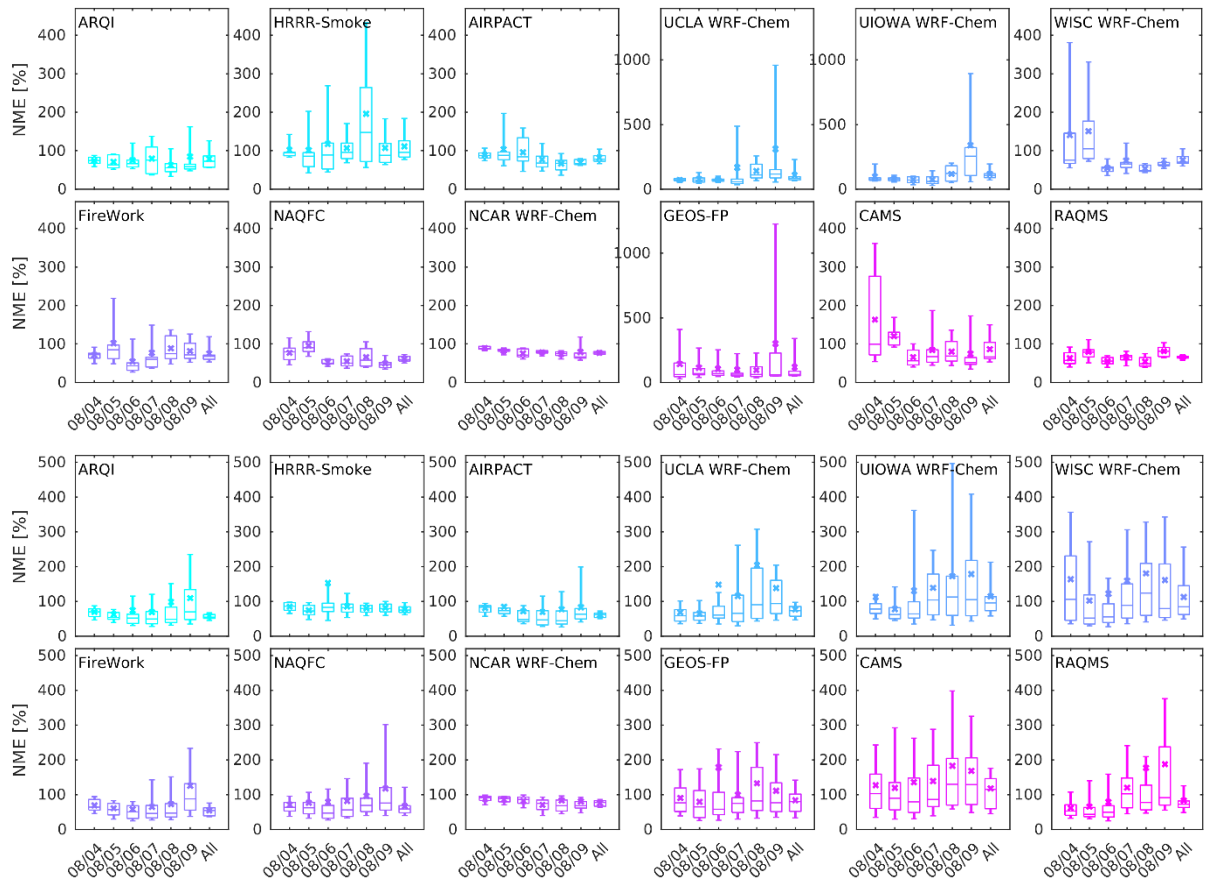
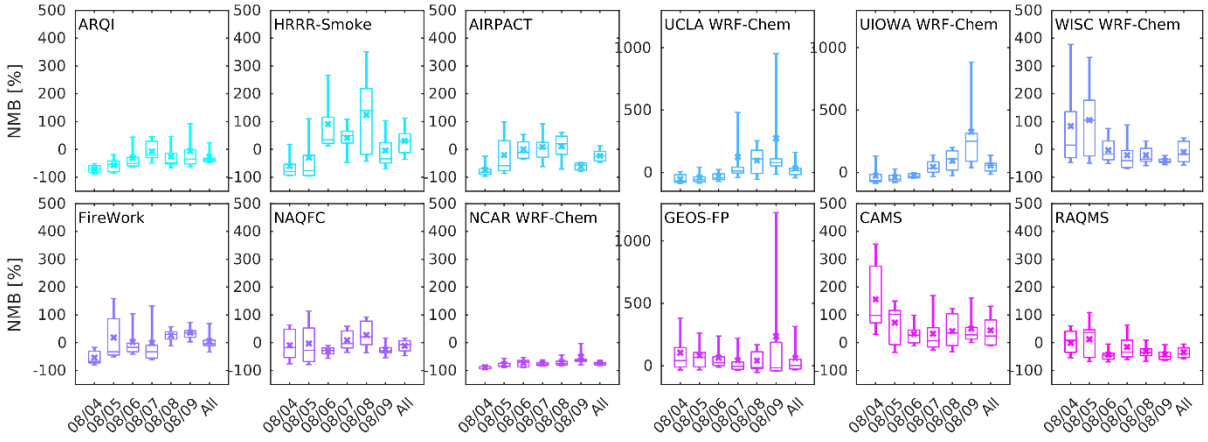


Figure S28. Boxplots of NME of PM2.5 enhancements for the two categories of surface sites: (a) fresh-plume (Fp) sites, (b) other sites (Ot), during 4-9 August 2019. Day-by-day results are shown with the comparison dates labeled for the x-axis in "month/day". The box labeled as "All" is for all data pairs of forecasts and observations. For each box, the central mark represents the median, and the bottom and top edges of box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the 10th and 90th percentiles. The crosses represent the average value.



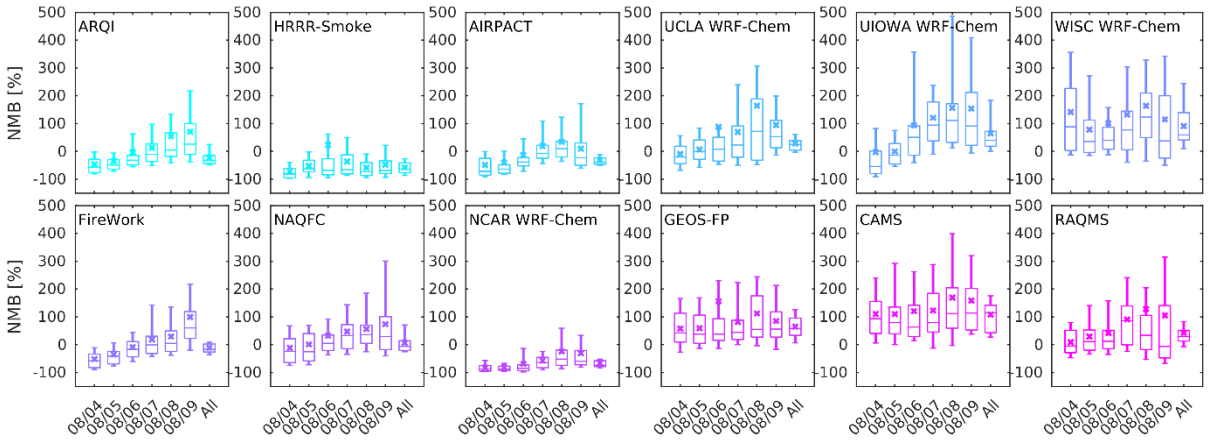


Figure S29. Similar as Figure S28, but for NMB.

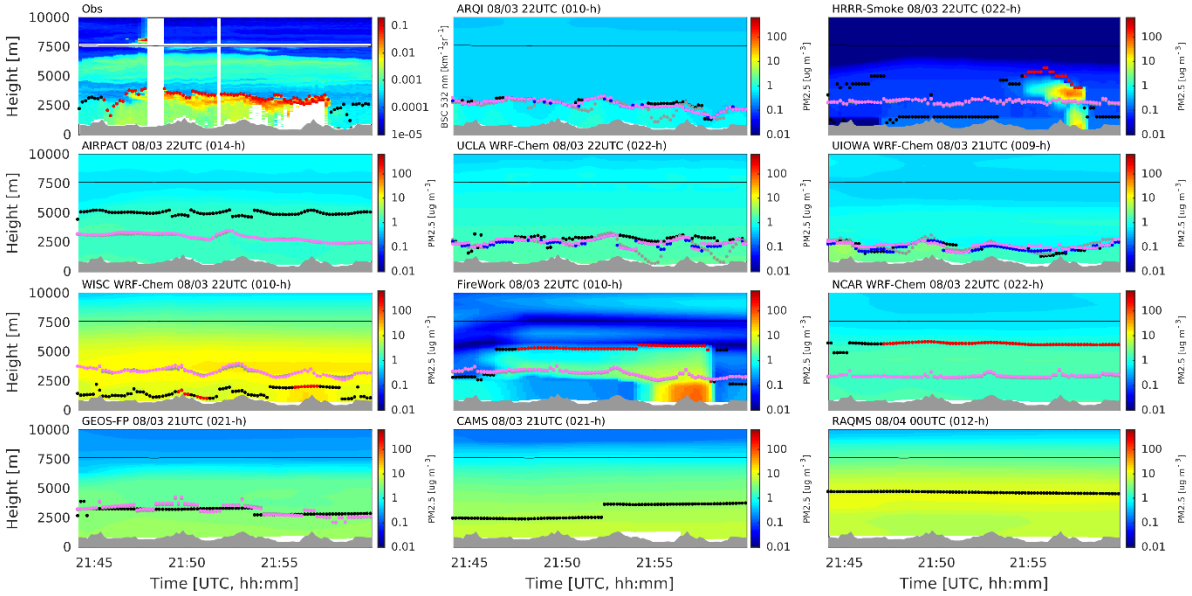


Figure S30. Comparison of vertical smoke structure based on the DIAL-HSRL observations along the transect D3T1 on 7 August (see details in Table 4 of the main text) through the smoke plume from the Williams Flats fire. Backscatter coefficient (532 nm) (top left panel) and PM2.5 concentrations forecasted by different models (other panels) are shown in log scale. The red dots are plume heights determined using the observed backscatter profiles or modeled PM2.5 mass concentration profiles. The black dots are the heights determined by the same method as the plume heights but using profiles sampled out of smoke plume, thus they can represent the mixed layer heights. The PBL heights for ARQI, UCLA WRF-Chem, and UIOWA WRF-Chem (the blue dots) are determined by using the virtual potential temperature profiles. For reference, the model diagnosed PBL heights are shown by the grey dots. For the other models, the PBL heights provided by the model data (the grey dots) are used. The pink dots are PBL heights for models at the closest output hour to 16:00 PDT, which are used to represent the maximum heights of mixed layer during the daytime. The black solid line shows flight heights.

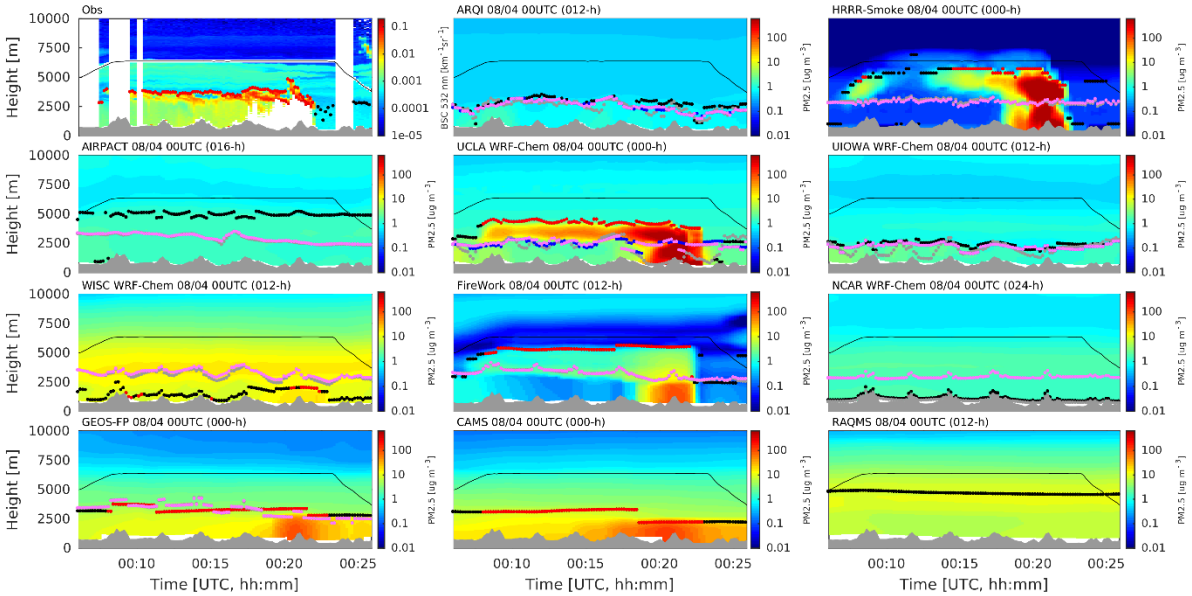


Figure S31. Similar to Figure S30, but for the fresh-plume transect D3T2.

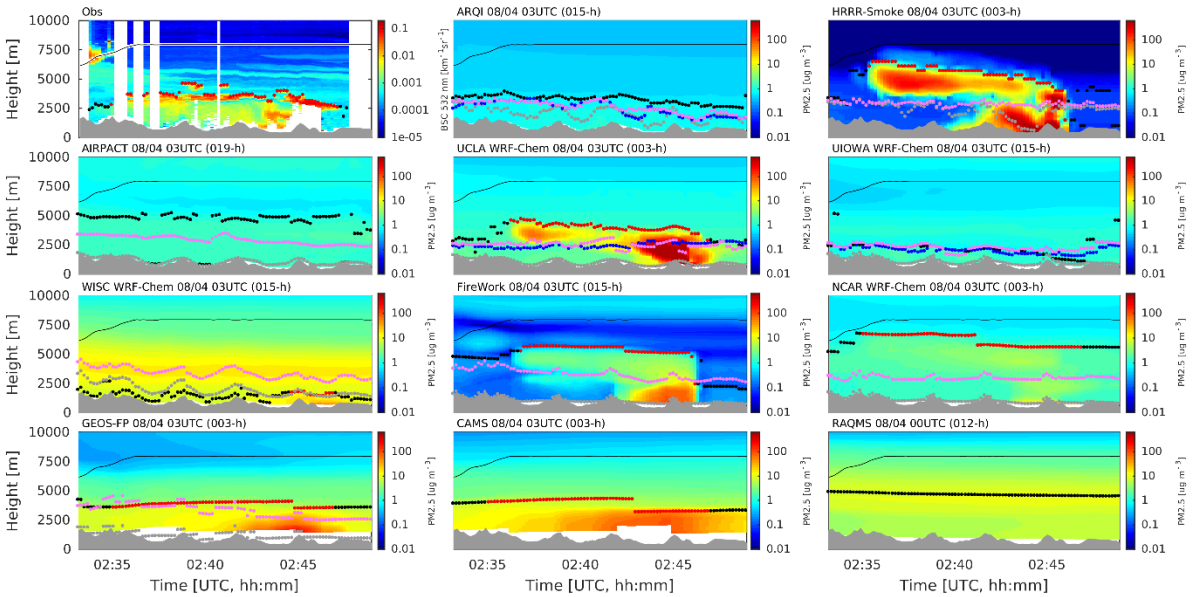


Figure S32. Similar to Figure S30, but for the fresh-plume transect D3T3.

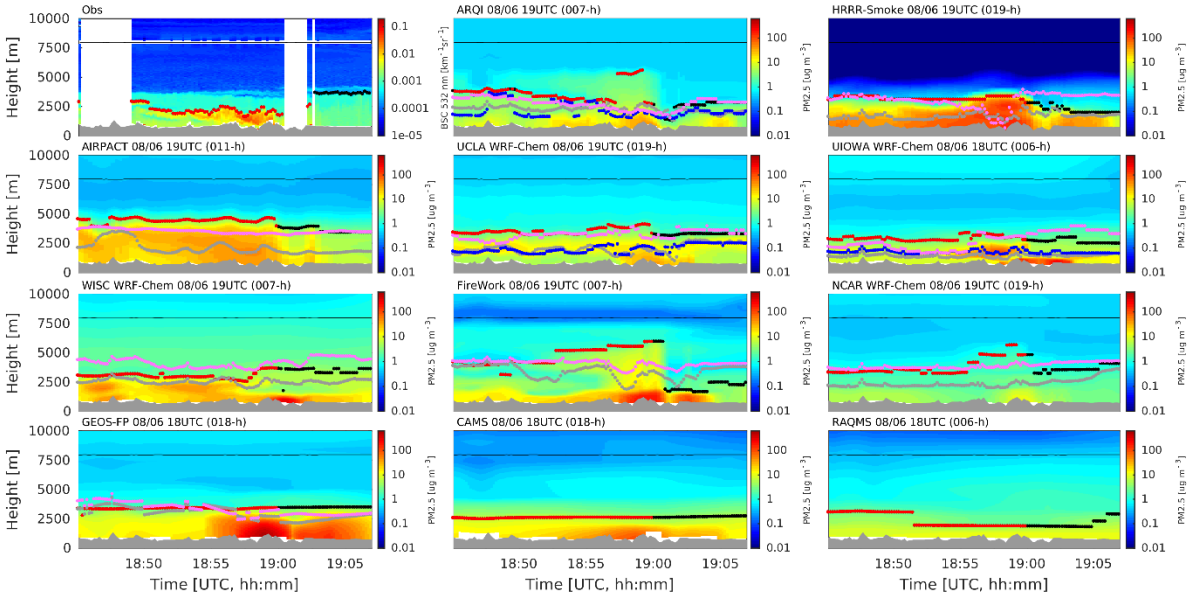


Figure S33. Similar to Figure S30, but for the fresh-plume transect D6T1.

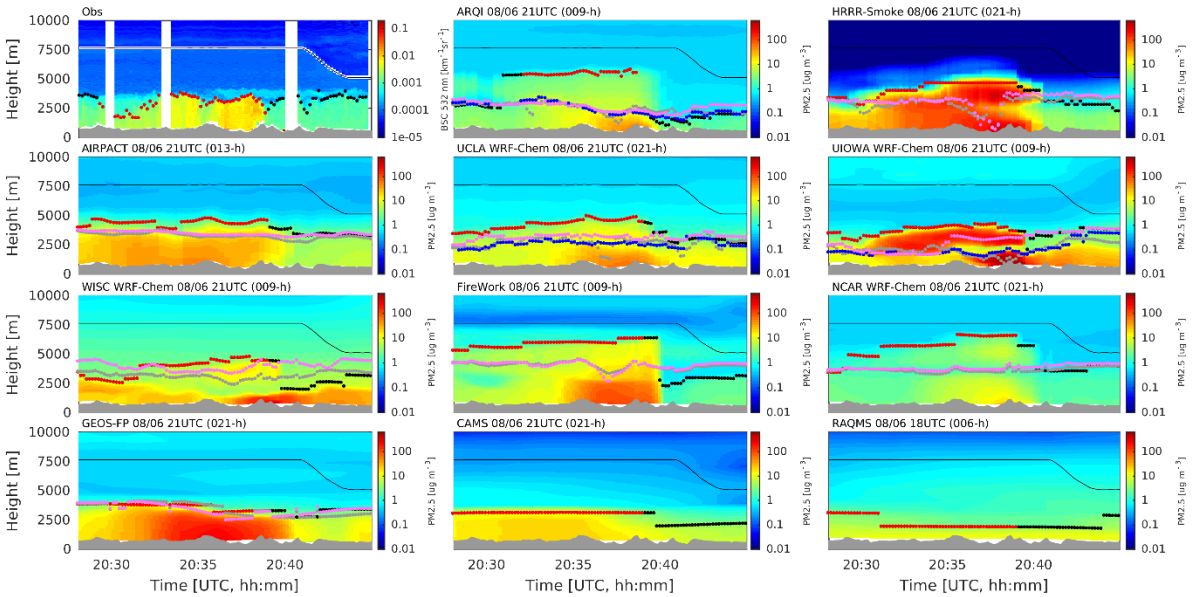


Figure S34. Similar to Figure S30, but for the fresh-plume transect D6T2.

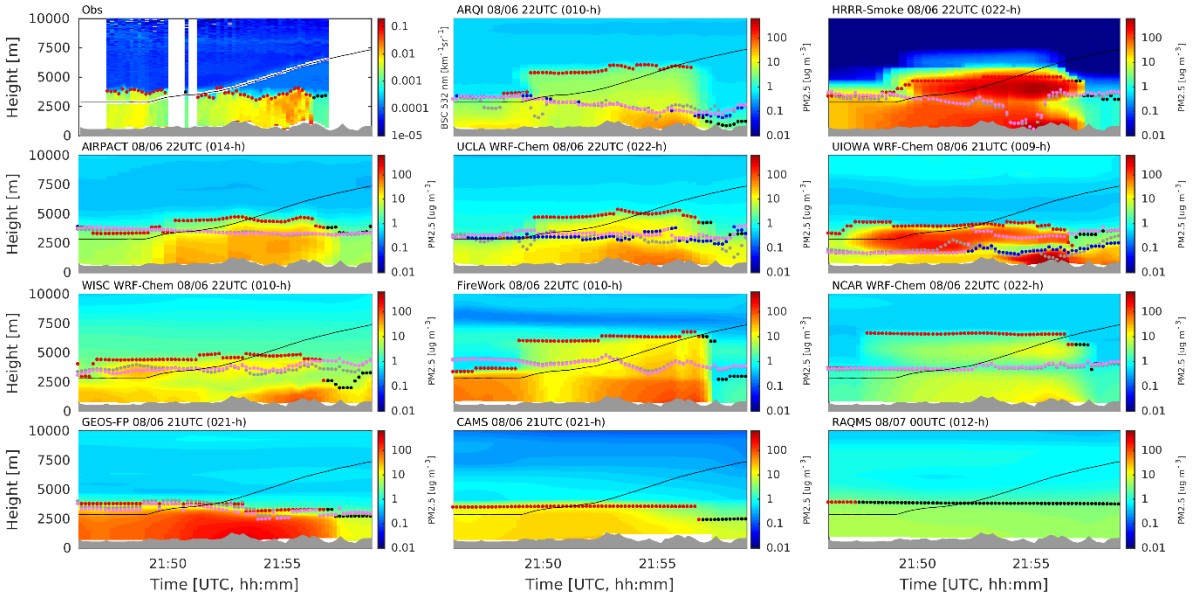


Figure S35. Similar to Figure S30, but for the fresh-plume transect D6T3.

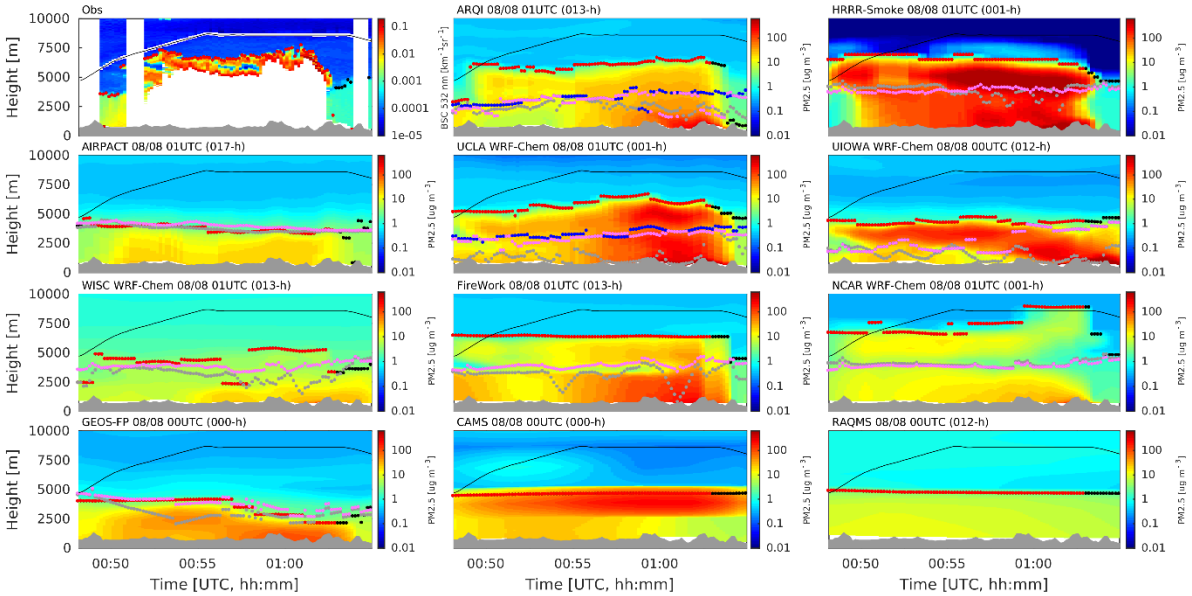


Figure S36. Similar to Figure S30, but for the fresh-plume transect D7T4.

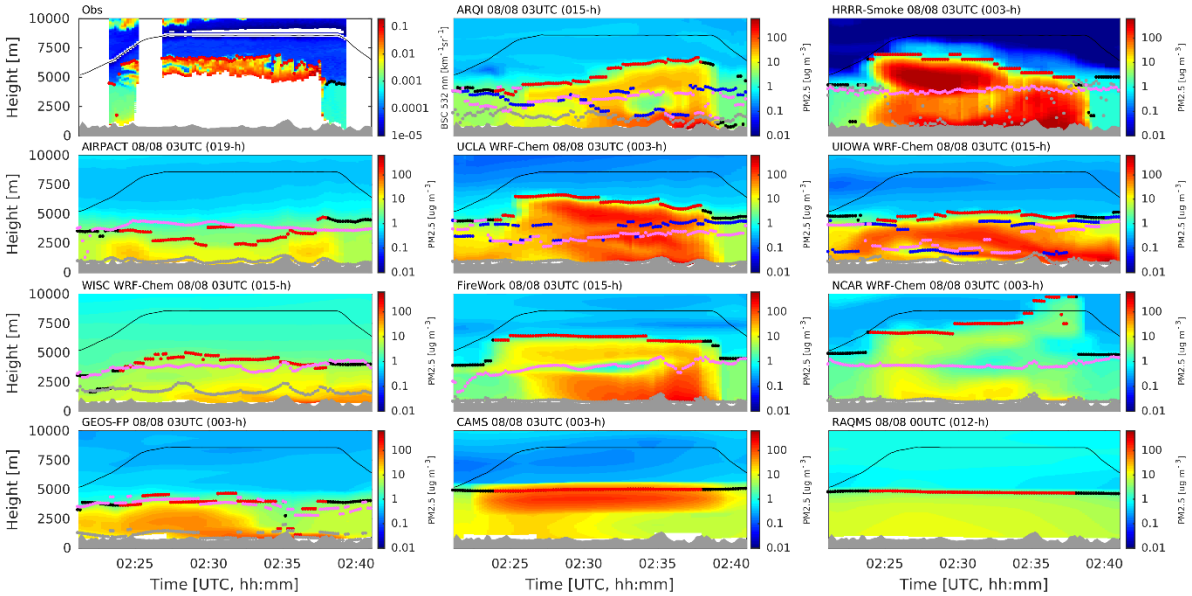


Figure S37. Similar to Figure S30, but for the fresh-plume transect D7T5.

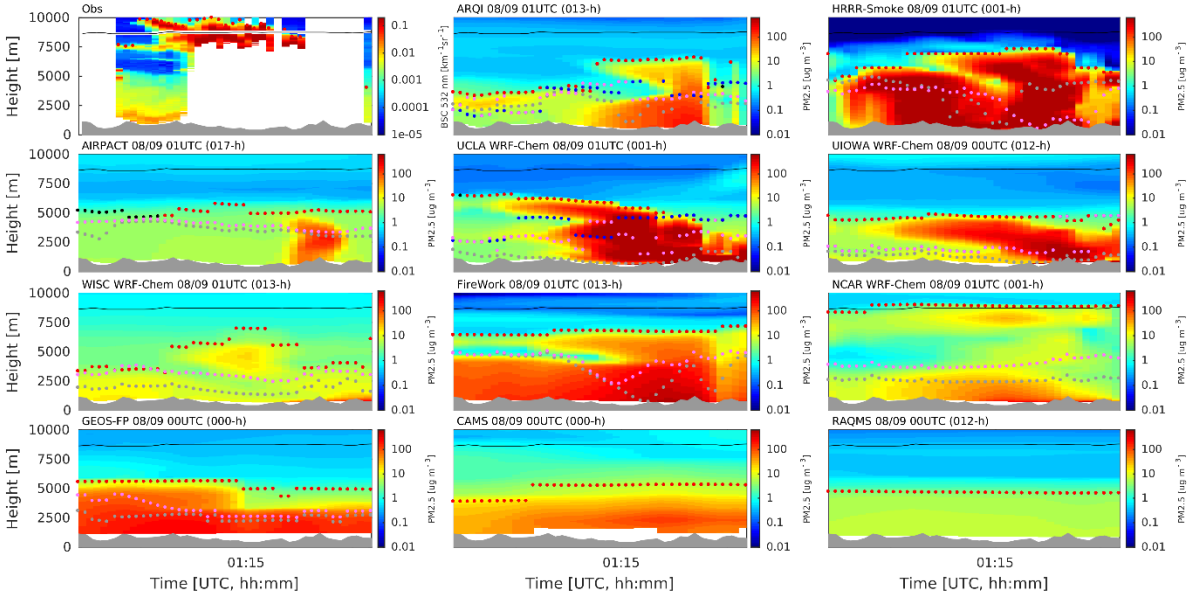


Figure S38. Similar to Figure S30, but for the fresh-plume transect D8T3.

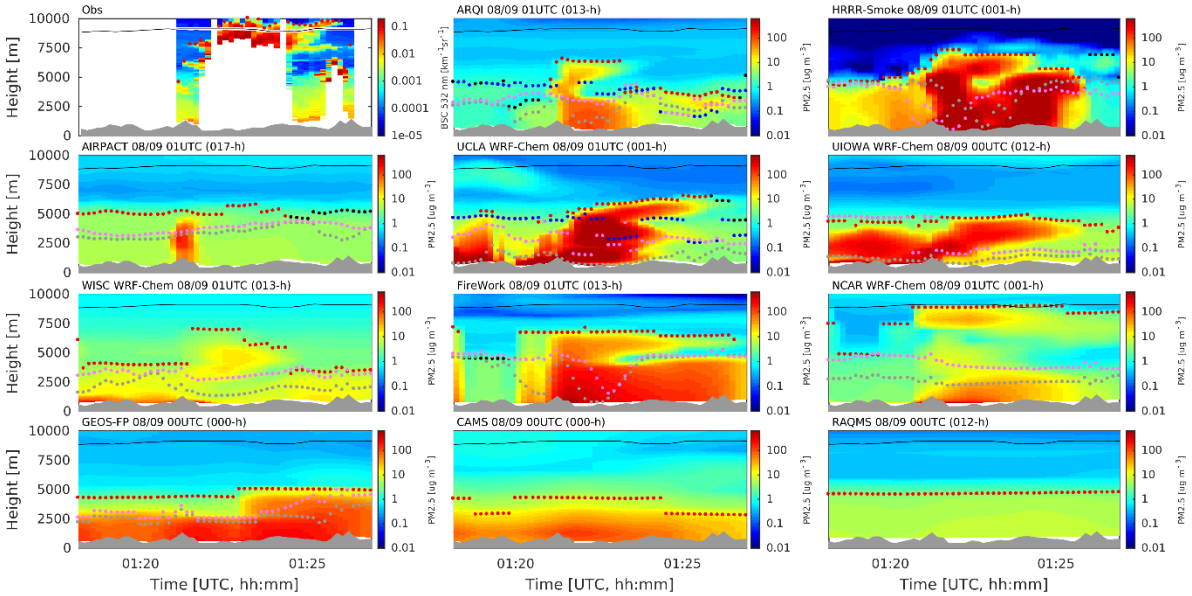


Figure S39. Similar to Figure S30, but for the fresh-plume transect D8T4.

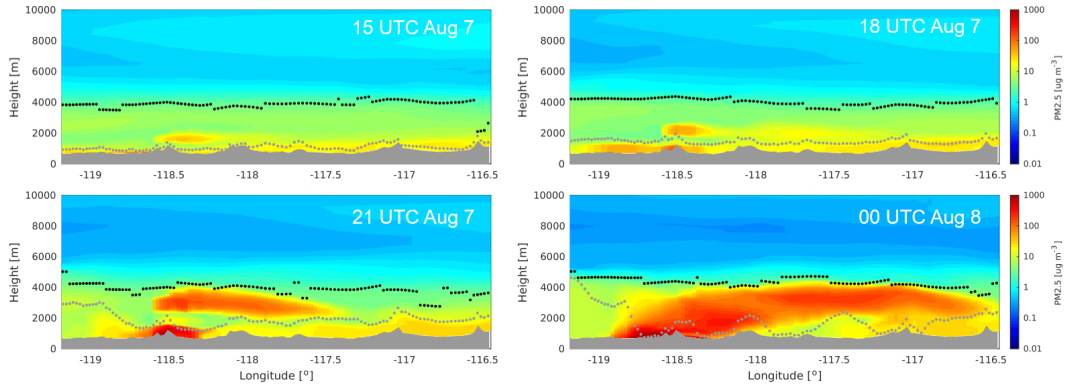


Figure S40. Vertical cross-sections of modelled PM2.5 concentrations by UIOWA WRF-Chem along the locations of flight transect D7T3. The valid times are labeled in each panel.

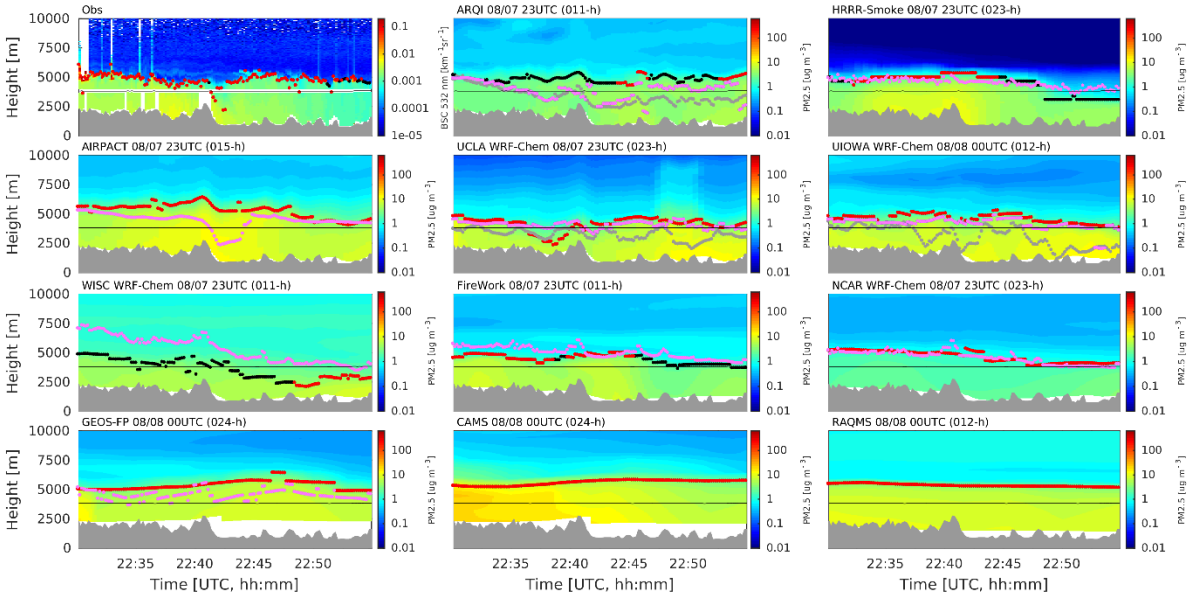


Figure S41. Similar to Figure S30, but for the aged-plume transect D7T2.

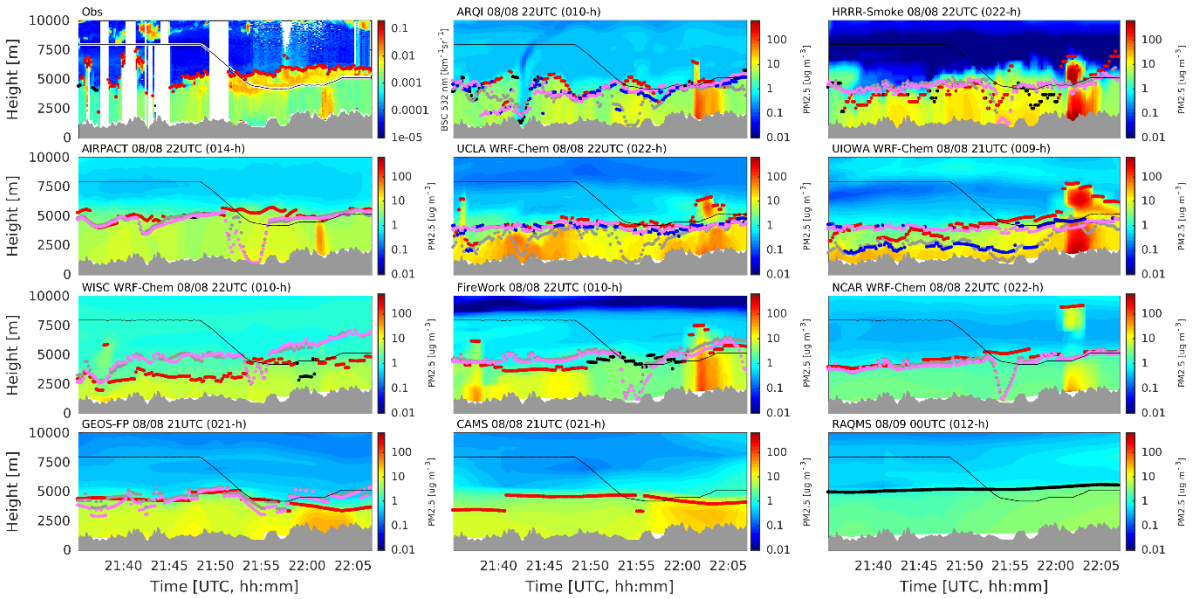


Figure S42. Similar to Figure S30, but for the aged plume transect D8T1.

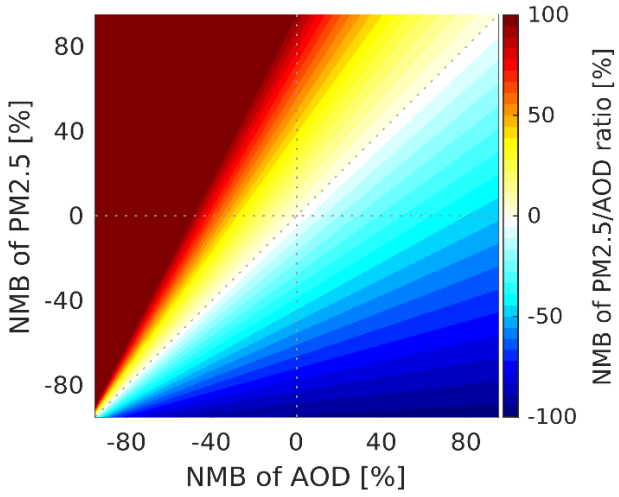


Figure S43. Ideal relationship of normalized mean bias (NMB) of surface PM2.5/AOD ratio against different combinations of NMB for AOD and surface PM2.5.